

GenCore version 5.1.9
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: June 12, 2006, 12:43:40 ; Search time 200 Seconds
(without alignments)
1682.556 Million cell updates/sec

Title: US-10-696-261-13

Perfect score: 3989
Sequence: 1 MADGYLPWMLDNLSEGR.....NNGLYTEPRPIGRVYLRPL 736

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 2589679 seqs, 457216429 residues

Total number of hits satisfying chosen parameters: 2589679

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

A_Geneseq_8: *
1: geneseqp1980s: *
2: geneseqp1990s: *
3: geneseqp2000s: *
4: geneseqp2001s: *
5: geneseqp2002s: *
6: geneseqp2003s: *
7: geneseqp2003bs: *
8: geneseqp2004s: *
9: geneseqp2005s: *
10: geneseqp2006s: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	3989	100.0	736	AAV71167	AAV71167 Adeno-sss
2	3989	100.0	736	ABBB80232	ABBB80232 AAV1 VP1
3	3989	100.0	736	ABR62762	ABR62762 Adeno ass
4	3989	100.0	736	ADZ76565	ADZ76565 Adeno-sss
5	3989	100.0	736	ADV70294	ADV70294 Primate a
6	3989	100.0	736	ADZ27069	ADZ27069 Adeno-sss
7	3989	100.0	736	ADZ27012	ADZ27012 Adeno-sss
8	3985	99.9	735	ADV67506	ADV67506 Amino aci
9	3968	99.5	736	ADZ27007	ADZ27007 Adeno-sss
10	3963	99.3	736	AAAB59847	AAAB59847 AAV6 caps
11	3963	99.3	736	ADZ76566	ADZ76566 Adeno-sss
12	3963	99.3	736	ADV70293	ADV70293 Primate a
13	3963	99.3	736	ADZ27070	ADZ27070 Adeno-sss
14	3897.5	97.7	737	ADZ27086	ADZ27086 Adeno-sss
15	3883.5	97.4	737	ADZ27010	ADZ27010 Adeno-sss
16	3883	97.3	736	ADZ27074	ADZ27074 Adeno-sss
17	3873	97.1	736	ADZ27008	ADZ27008 Adeno-sss
18	3853	96.6	736	ADZ27009	ADZ27009 Adeno-sss
19	3511	88.0	736	AAAB59846	AAAB59846 AAV3B cap
20	3494	87.6	736	AAAB59845	AAAB59845 AAV3A cap
21	3494	87.6	736	ABBB80233	ABBB80233 AAV3 VP1
22	3494	87.6	736	ABR62763	ABR62763 Adeno ass
23	3494	87.6	736	ADZ76572	ADZ76572 Adeno-sss

24	3494	87.6	736	9	ADV67507	ADV67507 Amino aci
25	3494	87.6	736	9	ADZ27067	ADZ27067 Adeno-sss
26	3483.5	87.3	737	9	ADZ26959	ADZ26959 Adeno-sss
27	3476	87.1	736	9	ADZ27085	ADZ27085 Adeno-sss
28	3473	87.1	736	7	ADZ76598	ADZ76598 Adeno-sss
29	3469	87.0	738	9	ADZ26960	ADZ26960 Adeno-sss
30	3467	86.9	736	6	ABBB80229	ABBB80229 AAV9 cap
31	3467	86.9	736	7	ABR62764	ABR62764 Adeno ass
32	3467	86.9	736	7	ADZ76601	ADZ76601 Adeno-sss
33	3467	86.9	736	9	ADZ67508	ADZ67508 Amino aci
34	3464.5	86.9	737	9	ADZ27080	ADZ27080 Adeno-sss
35	3464	86.8	736	7	ADZ76597	ADZ76597 Adeno-sss
36	3460	86.7	736	7	ADZ76599	ADZ76599 Adeno-sss
37	3458.5	86.7	737	9	ADZ26963	ADZ26963 Adeno-sss
38	3457	86.7	736	7	ADZ76600	ADZ76600 Adeno-sss
39	3456.5	86.7	735	9	ADZ27034	ADZ27034 Adeno-sss
40	3456.5	86.7	737	9	ADZ26966	ADZ26966 Adeno-sss
41	3456.5	86.7	737	9	ADZ26962	ADZ26962 Adeno-sss
42	3456	86.6	738	9	ADZ26951	ADZ26951 Adeno-sss
43	3455.5	86.6	737	9	ADZ26965	ADZ26965 Adeno-sss
44	3453.5	86.6	735	9	ADZ26996	ADZ26996 Adeno-sss
45	3453.5	86.6	735	9	ADZ26992	ADZ26992 Adeno-sss

ALIGNMENTS

RESULT 1
ID AAV71167 standard; protein; 736 AA.
AAV71167
XX
AC AAV71167;
XX
DT 08-SEP-2000 (first entry)
XX
DE Adeno-associated virus serotype 1 capsid protein VP1.
XX
KW Adeno-associated virus serotype 1; AAV-1; rep protein; capsid protein; cap protein; recombinant viral vector; gene delivery; gene therapy;
KW vaccine; transgene; VP1.
XX
OS Adeno-associated virus 1.
XX
PN W0200028061-A2.
XX
PD 18-MAY-2000.
XX
PF 02-NOV-1999; 99MO-US025694.
XX
PR 05-NOV-1998; 98US-0107114P.
XX
PA (UNP-) UNIV PENNSYLVANIA.
XX
PI Wilson JM, Xiao W;
XX
XX WPI; 2000-376571/32.
DR N-PSDB; AAD00772, AAD00777.
XX
PT Novel adeno-associated virus serotype 1 polynucleotide useful for preparation of medicament for delivery of a transgene to a host.
XX
PS Claim 7; Page 87-90; 108pp; English.
XX
CC The patent discloses an adeno-associated virus serotype 1 (AAV-1) DNA which is characterised by two inverted terminal repeats (ITR) and open reading frames for rep and capsid (cap) proteins. The rep reading frame encodes four proteins, Rep 78, Rep 68, Rep 52 and Rep 40, while the cap reading frame encodes three structural proteins, VP1, VP2 and VP3. The AAV-1 sequence or its fragments particularly ITRs, rep and cap coding regions, are useful in production of recombinant viral vectors for gene delivery. These vectors can be used as gene therapy vectors, vaccine vectors or antisense delivery vectors. The AAV-1 does not induce the formation of neutralising antibodies specific to any serotype of AAV

CC hence is useful for transforming host cells, and in the preparation of a
 CC medicament for the delivery of transgene to a host. The present sequence
 CC is an AAV-1 cap protein VP1 which is useful in the production of
 CC recombinant viral vector for gene delivery

XX Sequence 736 AA:

Query Match 100.0%; Score 3989; DB 3; Length 736;
 Best Local Similarity 100.0%; Pred. No. 5.2e-311;
 Matches 736; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MAADGTLPMLENTLSGIREMMDLKPAPKPRANQKODGGLVLPGYKYLGFENGLD 60
 DB 1 MAADGTLPMLENTLSGIREMMDLKPAPKPRANQKODGGLVLPGYKYLGFENGLD 60
 QY 61 KGEPVNAADAAALEHDKAYDQOLKAGDNPYLRYNHADAERFOERLOEDTSFGNUGRAVFO 120
 DB 61 KGEPVNAADAAALEHDKAYDQOLKAGDNPYLRYNHADAERFOERLOEDTSFGNUGRAVFO 120
 QY 121 AKRRVLEPLGLVEGAKTAPGKRPVEQSPQEPDSSSGIGKTGQOPAKKRLNFGQTGDS 180
 DB 121 AKRRVLEPLGLVEGAKTAPGKRPVEQSPQEPDSSSGIGKTGQOPAKKRLNFGQTGDS 180
 QY 181 SVDPDQPLGEPPTPAVGPPTMASGGGAPMADNNEGADVGNAAGNWHCDSTWLGDRVI 240
 DB 181 SVDPDQPLGEPPTPAVGPPTMASGGGAPMADNNEGADVGNAAGNWHCDSTWLGDRVI 240
 QY 241 TTSRTWALPTVNNHLYKQISSASTGASNDNHFGYSTPMPGYDFENRFGHCFSPRDMQRL 300
 DB 241 TTSRTWALPTVNNHLYKQISSASTGASNDNHFGYSTPMPGYDFENRFGHCFSPRDMQRL 300
 QY 301 INNMGFRPKRLNPKLFNIOVKVTTNDGYTTIANNTSTVOYFSDSEYOLPYVLSAQH 360
 DB 301 INNMGFRPKRLNPKLFNIOVKVTTNDGYTTIANNTSTVOYFSDSEYOLPYVLSAQH 360
 QY 361 GCLPPPADYFEMIPOYGYLTLLNNGSOAVGRSFFCLEYFSPSOMLRTGNNTFFSYTFEEVP 420
 DB 361 GCLPPPADYFEMIPOYGYLTLLNNGSOAVGRSFFCLEYFSPSOMLRTGNNTFFSYTFEEVP 420
 QY 421 FHSSYAHOSGLDRLNMPLDIOYLYLNRTONOGSAQNKDLLFSRGSFAGMSVOPKMWLP 480
 DB 421 FHSSYAHOSGLDRLNMPLDIOYLYLNRTONOGSAQNKDLLFSRGSFAGMSVOPKMWLP 480
 QY 481 GPCYRQORVSKTTDNNNSNFTWTGASKYNLNGRESIINPGTMAASHKDEDEKFFPMGSV 540
 DB 481 GPCYRQORVSKTTDNNNSNFTWTGASKYNLNGRESIINPGTMAASHKDEDEKFFPMGSV 540
 QY 541 MIFGKSAGASNTALDNVMTDEBEIKATNPVATERFGTVAVNFOSSSTDPAIGDVAMG 600
 DB 541 MIFGKSAGASNTALDNVMTDEBEIKATNPVATERFGTVAVNFOSSSTDPAIGDVAMG 600
 QY 601 ALPGMWQDRDYYLQGITWAKIPHTDGHFHPSPLMGFGKLPQILIKNTVPANPPA 660
 DB 601 ALPGMWQDRDYYLQGITWAKIPHTDGHFHPSPLMGFGKLPQILIKNTVPANPPA 660
 QY 661 EBSATKFASTITTOYSTQGVSEIEMELQKENSKRANPEVOYTSNYAASAVDFTVDNGL 720
 DB 661 EBSATKFASTITTOYSTQGVSEIEMELQKENSKRANPEVOYTSNYAASAVDFTVDNGL 720
 QY 721 YTEPRPIGRYTLRPL 736
 DB 721 YTEPRPIGRYTLRPL 736

RESULT 2
 ABB80232
 ID ABB80232 standard; protein: 736 AA.

XX ABB80232;
 XX 20-NOV-2003 (first entry)
 XX AAV1 vp1 protein.

XX Adeno-associated virus; AAV; serotype 9; rep; cap; vp1; vp2; vp3;
 KW splice variant; transgene.
 XX
 XX Adeno associated virus serotype 1.
 OS
 XX
 PN W02003052052-AA2.

PD 26-JUN-2003.

PF 12-NOV-2002; 2002WO-US033631.

PR 17-DEC-2001; 2001US-0341150P.

PR 05-JUN-2002; 2002US-0386132P.

PA (UYPE-) UNIV PENNSYLVANIA.

PI Gao G, Wilson JM, Alvira M;

XX WPI, 2003-523523/49.

PT New isolated adeno-associated virus (AAV) comprising an AAV9 capsid,
 useful for preparing a medicament for delivering a transgene to a cell.

PS Disclosure; Fig 2; 76pp; English.

XX The sequences given in ABB80231-34 represent vp1 proteins derived from
 CC various adeno-associated virus (AAV) serotypes. These sequences were used
 CC in the scope of the invention for comparison with the cap protein derived
 CC from AAV serotype 9. The AAV capsid comprises three proteins VP1, VP2 and
 CC VP3, which are alternative splice variants. The AAV or the nucleic acid
 CC molecule is useful for preparing a medicament for delivering a transgene
 CC to a cell

XX Sequence 736 AA:

Query Match 100.0%; Score 3989; DB 6; Length 736;
 Best Local Similarity 100.0%; Pred. No. 5.2e-311;
 Matches 736; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MAADGTLPMLENTLSGIREMMDLKPAPKPRANQKODGGLVLPGYKYLGFENGLD 60
 DB 1 MAADGTLPMLENTLSGIREMMDLKPAPKPRANQKODGGLVLPGYKYLGFENGLD 60
 QY 61 KGEPVNAADAAALEHDKAYDQOLKAGDNPYLRYNHADAERFOERLOEDTSFGNUGRAVFO 120
 DB 61 KGEPVNAADAAALEHDKAYDQOLKAGDNPYLRYNHADAERFOERLOEDTSFGNUGRAVFO 120
 QY 121 AKRRVLEPLGLVEGAKTAPGKRPVEQSPQEPDSSSGIGKTGQOPAKKRLNFGQTGDS 180
 DB 121 AKRRVLEPLGLVEGAKTAPGKRPVEQSPQEPDSSSGIGKTGQOPAKKRLNFGQTGDS 180
 QY 181 SVDPDQPLGEPPTPAVGPPTMASGGGAPMADNNEGADVGNAAGNWHCDSTWLGDRVI 240
 DB 181 SVDPDQPLGEPPTPAVGPPTMASGGGAPMADNNEGADVGNAAGNWHCDSTWLGDRVI 240
 QY 241 TTSRTWALPTVNNHLYKQISSASTGASNDNHFGYSTPMPGYDFENRFGHCFSPRDMQRL 300
 DB 241 TTSRTWALPTVNNHLYKQISSASTGASNDNHFGYSTPMPGYDFENRFGHCFSPRDMQRL 300
 QY 301 INNMGFRPKRLNPKLFNIOVKVTTNDGYTTIANNTSTVOYFSDSEYOLPYVLSAQH 360
 DB 301 INNMGFRPKRLNPKLFNIOVKVTTNDGYTTIANNTSTVOYFSDSEYOLPYVLSAQH 360
 QY 361 GCLPPPADYFEMIPOYGYLTLLNNGSOAVGRSFFCLEYFSPSOMLRTGNNTFFSYTFEEVP 420
 DB 361 GCLPPPADYFEMIPOYGYLTLLNNGSOAVGRSFFCLEYFSPSOMLRTGNNTFFSYTFEEVP 420
 QY 421 FHSSYAHOSGLDRLNMPLDIOYLYLNRTONOGSAQNKDLLFSRGSFAGMSVOPKMWLP 480
 DB 421 FHSSYAHOSGLDRLNMPLDIOYLYLNRTONOGSAQNKDLLFSRGSFAGMSVOPKMWLP 480
 QY 481 GPCYRQORVSKTTDNNNSNFTWTGASKYNLNGRESIINPGTMAASHKDEDEKFFPMGSV 540

```
|||||
Db      481 GPCRRQGRVSKTKTDNNNSNFTWTGASKYLNNGRESIIINPTAASHKDEDEKFFPMGSV 540
Qy      541 MIFKESAGASNTALDWMITDEBEIKATNPVATERFGTVAVNFQSSSTDPATGDVHAMG 600
Db      541 MIFKESAGASNTALDWMITDEBEIKATNPVATERFGTVAVNFQSSSTDPATGDVHAMG 600
Qy      601 ALPGMWQDRDVLVYLGQIPMAKIPHTDGHFHPSPLMGFGGLKNPPQILIKNTVPANPPA 660
Db      601 ALPGMWQDRDVLVYLGQIPMAKIPHTDGHFHPSPLMGFGGLKNPPQILIKNTVPANPPA 660
Qy      661 EFSATKFASTITQVSTGQVSEIEMELQKENSKRMPNEVOYTSNYAKSANVDFTVDNGL 720
Db      661 EFSATKFASTITQVSTGQVSEIEMELQKENSKRMPNEVOYTSNYAKSANVDFTVDNGL 720
Qy      721 YTEBRPIGTRYLTRPL 736
Db      721 YTEBRPIGTRYLTRPL 736
```

RESULT 3
ABR62762
ID ABR62762 standard; protein; 736 AA.

AC ABR62762;

DT 06-NOV-2003 (first entry)

XX Adeno associated virus 1 capsid protein vp1.

KM AAV; AAV1; capsid; vector; gene therapy; antisense therapy; vaccine.

OS Adeno associated virus.

PN MO2003052051-A2.

PD 26-JUN-2003.

PF 12-NOV-2002; 2002MO-US033630.

PR 17-DEC-2001; 2001US-0341151P.

PR 01-MAY-2002; 2002US-0377133P.

PR 05-JUN-2002; 2002US-0386122P.

PA (UYPE-) UNIV PENNSYLVANIA.

PI Gao G, Wilson JM, Alvirra M;

DR WPI; 2003-523522/49.

PT New adeno-associated virus (AAV) comprising an AAV capsid having an amino

PT acid sequence of AAV8, useful for preparing a medicament for delivery of

PT a transgene to a cell and for treating cystic fibrosis or hemophilia B.

PS Disclosure; Fig 2A-C; 82bp; English.

XX The present sequence is the protein sequence of the vp1 capsid protein of

CC adeno associated virus serotype 1 (AAV1). The invention provides the

CC nucleic acid and amino acid sequences of novel AAV8 and fragments of

CC these sequences. Each of these fragments may be used in a variety of

CC vector systems and host cells. Among the desirable fragments are the cap

CC proteins, including vp1, vp2, vp3 and hypervariable regions, the rep

CC encoding these proteins. The fragments may be used alone, in combination

CC with other AAV8 sequences or in combination with elements from other AAV

CC or non-AAV viral sequences in the production of recombinant AAV and for

CC use as antisense delivery vectors, gene therapy vectors or vaccine

CC vectors. A claimed molecule comprises a cap protein of a functional AAV

CC rep gene from a serotype selected from AAV1, AAV2, AAV3, AAV4, AAV5 and

CC AAV6

XX Sequence 736 AA;

Query Match 100.0%; Score 3989; DB 7; Length 736;
Best Local Similarity 100.0%; Pred. No.5.2e-111;
Matches 736; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy      1 MAADGYLPDWLENDLSEGIRESMDLKPAGKPKANQOKODGRGLVPGYKYLGFENGLD 60
Db      1 MAADGYLPDWLENDLSEGIRESMDLKPAGKPKANQOKODGRGLVPGYKYLGFENGLD 60
Qy      61 KGEPPNADAAALAEHDKAYVQOLKAGNPYLRYNHADAERQERLOEPTSGGNIGRAVFQ 120
Db      61 KGEPPNADAAALAEHDKAYVQOLKAGNPYLRYNHADAERQERLOEPTSGGNIGRAVFQ 120
Qy      121 AKKRVLEPLGLVEGATAPGKKRPVQSPQEPSSSGIGTGQPAKKRLNFQGTQDSE 180
Db      121 AKKRVLEPLGLVEGATAPGKKRPVQSPQEPSSSGIGTGQPAKKRLNFQGTQDSE 180
Qy      181 SVDPPOPLGEPPTAPPAVGPPTMASGCGAPMADNEGADGVNAGNMHCDSTWLAGRVI 240
Db      181 SVDPPOPLGEPPTAPPAVGPPTMASGCGAPMADNEGADGVNAGNMHCDSTWLAGRVI 240
Qy      241 TTSRTWALPTYNHLYKQISSASTGASNDNHFGYSTPGYDFNRFCHFSRDMQRL 300
Db      241 TTSRTWALPTYNHLYKQISSASTGASNDNHFGYSTPGYDFNRFCHFSRDMQRL 300
Qy      301 INNMGFRPKRLNPKLFNIQVKEVTNDGYTTIANNTSTVQVPSDSEYOLPYLGSAAQ 360
Db      301 INNMGFRPKRLNPKLFNIQVKEVTNDGYTTIANNTSTVQVPSDSEYOLPYLGSAAQ 360
Qy      361 GCLPPFPADYFMIPOYGYLTINNGSQAVGRSSFYCLEYFPSOMLRTGNFTFSYTFEEVP 420
Db      361 GCLPPFPADYFMIPOYGYLTINNGSQAVGRSSFYCLEYFPSOMLRTGNFTFSYTFEEVP 420
Qy      421 FHSSYAHSGSLDRIMNPLIDQYLYINRTQNGSQAQNKDLFSGSPAGMSVQPKWLP 480
Db      421 FHSSYAHSGSLDRIMNPLIDQYLYINRTQNGSQAQNKDLFSGSPAGMSVQPKWLP 480
Qy      481 GPCYRQGRVSKTKTDNNNSNFTWTGASKYLNNGRESIIINPTAASHKDEDEKFFPMGSV 540
Db      481 GPCYRQGRVSKTKTDNNNSNFTWTGASKYLNNGRESIIINPTAASHKDEDEKFFPMGSV 540
Qy      541 MIFKESAGASNTALDWMITDEBEIKATNPVATERFGTVAVNFQSSSTDPATGDVHAMG 600
Db      541 MIFKESAGASNTALDWMITDEBEIKATNPVATERFGTVAVNFQSSSTDPATGDVHAMG 600
Qy      601 ALPGMWQDRDVLVYLGQIPMAKIPHTDGHFHPSPLMGFGGLKNPPQILIKNTVPANPPA 660
Db      601 ALPGMWQDRDVLVYLGQIPMAKIPHTDGHFHPSPLMGFGGLKNPPQILIKNTVPANPPA 660
Qy      661 EFSATKFASTITQVSTGQVSEIEMELQKENSKRMPNEVOYTSNYAKSANVDFTVDNGL 720
Db      661 EFSATKFASTITQVSTGQVSEIEMELQKENSKRMPNEVOYTSNYAKSANVDFTVDNGL 720
Qy      721 YTEBRPIGTRYLTRPL 736
Db      721 YTEBRPIGTRYLTRPL 736
```

RESULT 4
ADE76565

ID ADE76565 standard; protein; 736 AA.

AC ADE76565;

DT 29-JAN-2004 (first entry)

DE Adeno-associated virus (AAV) related protein, SEQ ID No 64.

XX adeno-associated virus; AAV; cytosolic; antiproliferative; antineoplastic;

KM antidiabetic; antidiabetic; antidiabetic; antidiabetic;

KM dermatological; antiinflammatory; gene therapy; vaccine;

KM rheumatoid arthritis; multiple sclerosis; diabetes;

KM autoimmune thyroiditis; scleroderma; Crohn's disease.

XX Unidentified.
OS
XX
PN EP1310571-A2.
PD 14-MAY-2003.
PF 12-NOV-2002; 2002EP-00257826.
XX
XX 13-NOV-2001; 2001US-0350607P.
PR 17-DEC-2001; 2001US-0341117P.
PR 01-MAY-2002; 2002US-0377066P.
PR 05-JUN-2002; 2002US-0386675P.
XX
PA (UYPE-) UNIV PENNSYLVANIA.
PI Gao G, Wilson JM, Alvira M;
XX
XX WPI; 2003-450984/43.
DR
XX Detecting adeno-associated virus sequences in a sample, useful for e.g.
PT preventing or treating hyperproliferative or autoimmune diseases,
PT comprises subjecting a sample having a DNA to amplification via
PT polymerase chain reaction.
XX
XX Disclosure, SEQ ID NO 64; 419pp; English.
XX
XX The invention relates to a novel method for detecting adeno-associated
CC virus (AAV) sequences in a sample, which comprises subjecting a sample
CC containing a DNA to amplification via a polymerase chain reaction (PCR).
CC The AAV sequence have the following activities: cytostatic,
CC antidiabetic, antirheumatic, antiarthritic, neuroprotective,
CC antipsoriatic, antithyroid, dermatological, and antiinflammatory. The AAV
CC sequence can be used in gene therapy or as part of a vaccine to treat
CC disorders. The method is useful in detecting and/or identifying AAV
CC sequences and isolating novel sequences that are identified. The
CC sequences may be used e.g. for preventing or treating hyperproliferative
CC conditions such as cancers and psoriasis, and other autoimmune diseases
CC like rheumatoid arthritis, multiple sclerosis, diabetes, autoimmune
CC thyroiditis, scleroderma or Crohn's disease. This sequence represents an
CC AAV related protein sequence of the invention.
XX
XX Sequence 736 AA.
SQ
Query Match 100.0%; Score 3989; DB 7; Length 736;
Best Local Similarity 100.0%; Pred. No. 5, 2e-311;
Matches 736; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MAADGYIPDWLENDLSGIREMMDLKGPAPKPKANQKODGRGLVPGYKYLGPENGLD 60
DB 1 MAADGYIPDWLENDLSGIREMMDLKGPAPKPKANQKODGRGLVPGYKYLGPENGLD 60
QY 61 KGEFVNADAAALAEHDAVYDOOLKAGNPFYLRVNHADAEPQERLOEDTSFGNLRGAVFQ 120
DB 61 KGEFVNADAAALAEHDAVYDOOLKAGNPFYLRVNHADAEPQERLOEDTSFGNLRGAVFQ 120
QY 121 AKKRVLEPLGVBEAGTAPGKKRPVEQSPQEPDSSGIGTGOQPAKKRLNFGQTSDE 180
DB 121 AKKRVLEPLGVBEAGTAPGKKRPVEQSPQEPDSSGIGTGOQPAKKRLNFGQTSDE 180
QY 121 AKKRVLEPLGVBEAGTAPGKKRPVEQSPQEPDSSGIGTGOQPAKKRLNFGQTSDE 180
DB 121 AKKRVLEPLGVBEAGTAPGKKRPVEQSPQEPDSSGIGTGOQPAKKRLNFGQTSDE 180
QY 181 SVDPDPOGLGEPBPATPAVGPPTTMASSGGAPMADNNEGADVGNAAGMWHCDSTLGRVI 240
DB 181 SVDPDPOGLGEPBPATPAVGPPTTMASSGGAPMADNNEGADVGNAAGMWHCDSTLGRVI 240
QY 241 TTSRTTAAALPTVNNHLTKOISSASTGASNDNHFGYSTPMGYPFNFPHCFSPRDQRL 300
DB 241 TTSRTTAAALPTVNNHLTKOISSASTGASNDNHFGYSTPMGYPFNFPHCFSPRDQRL 300
QY 301 INNNWGRPRKRLNFKLNIQVKEVTTNDGVTIANNLSTVQVESDSEYQLPYVLGSAHQ 360
DB 301 INNNWGRPRKRLNFKLNIQVKEVTTNDGVTIANNLSTVQVESDSEYQLPYVLGSAHQ 360
QY 361 GCLPPPADVPMIQYGYLTILNNGSQAVGSSPYCLEYFSPQMLRTGNNPFTSTPEBVP 420

DB 361 GCLPPPADVPMIQYGYLTILNNGSQAVGSSPYCLEYFSPQMLRTGNNPFTSTPEBVP 420
QY 421 FHSSVYHQSGLDLRNLNPLIDQYLYINRTONQSGAQNODLFRSSPAGMSVQPKNWLP 480
DB 421 FHSSVYHQSGLDLRNLNPLIDQYLYINRTONQSGAQNODLFRSSPAGMSVQPKNWLP 480
QY 481 GPCYRQORVSKTKTDNNNSNFTWTGASKYNLNGRESIINPGTMAHKKDEDFPFMGSV 540
DB 481 GPCYRQORVSKTKTDNNNSNFTWTGASKYNLNGRESIINPGTMAHKKDEDFPFMGSV 540
QY 541 MIFGKESAGASNTALDNWMITDEEIKATNPATERFGTVAVNFQSSSTDPAQDVHAMG 600
DB 541 MIFGKESAGASNTALDNWMITDEEIKATNPATERFGTVAVNFQSSSTDPAQDVHAMG 600
QY 601 ALPGMTWODRDVYLQGPPIAKTIPHTDGHFHPSPLMGFGKLPPOILLIKNTEVPANPPA 660
DB 601 ALPGMTWODRDVYLQGPPIAKTIPHTDGHFHPSPLMGFGKLPPOILLIKNTEVPANPPA 660
QY 661 EFSATKFASTITQYSGOVSEIEMELQKENSRRKNPEVOYTSNYSKANVDPTVNNGL 720
DB 661 EFSATKFASTITQYSGOVSEIEMELQKENSRRKNPEVOYTSNYSKANVDPTVNNGL 720
QY 721 YTEPRPIGTRVYLRPL 736
DB 721 YTEPRPIGTRVYLRPL 736
RESULT 5
ADV70294
ID ADV70294 standard; protein; 736 AA.
AC ADV70294;
XX
DT 10-MAR-2005 (first entry)
XX
DE Primate adeno-associated virus 1 capsid protein VP1.
XX
XX immunosuppressive; gene therapy; immunogenicity; virus inactivation;
KW hemophilia; Pepck deficiency; galactosemia; phenylketonuria;
KW Crigler-Najjar syndrome; Gout; Zellweger syndrome; ischemia;
KW atherosclerosis; thrombosis; embolism; Parkinson's disease;
KW congestive heart failure; cancer; inflammation; immune disorder;
KW muscular dystrophy; diabetes; VP1.
XX
XX Adeno-associated virus 1.
OS
OS WO2004112727-A2.
XX
XX 29-DEC-2004.
PD
XX 21-JUN-2004; 2004WO-US019884.
PF
XX 19-JUN-2003; 2003US-0480395P.
PR 30-APR-2004; 2004US-0567310P.
PR 03-JUN-2004; 2004US-0576501P.
XX
XX (AVTG-) AVIGEN INC.
PA
XX
PI Abetman AB, Colosi P, Lochrie MA, Surosky RT;
XX
DR WPI; 2005-048755/05.
XX
XX New mutated adeno-associated virus (AAV) capsid protein that when present
PT in an AAV virion imparts decreased immunoreactivity to the virion as
PT compared to the corresponding wild-type virion, useful for treating e.g.
PT hemophilia.
XX
XX Example 5; SEQ ID NO 20; 136pp; English.
PS
XX The invention describes a mutated adeno-associated virus (AAV) capsid
CC protein that when present in an AAV virion imparts decreased
CC immunoreactivity to the virion as compared to the corresponding wild-type

CC virion. Also described are: a polynucleotide encoding the mutated protein
 CC above; a recombinant AAV virion comprising the mutated protein above; and
 CC delivering a recombinant AAV virion to a cell or tissue of a vertebrate
 CC subject. The recombinant AAV virion is useful for delivering a
 CC heterologous nucleic acid molecule to a cell or tissue of a vertebrate
 CC subject, where the protein encoded by the heterologous nucleic acid
 CC molecule is expressed at a level that provides a therapeutic effect,
 CC where the recombinant AAV virion may comprise a non-primate, mammalian
 CC AAV capsid protein that when present in an AAV virion imparts decreased
 CC immunoreactivity to the virion as compared to immunoreactivity of primate
 CC AAV-2, and the heterologous nucleic acid molecule, where the heterologous
 CC nucleic acid molecule encodes a therapeutic protein and is operably
 CC linked to control elements capable of directing the in vivo transcription
 CC and translation of the protein. The protein or the recombinant AAV virion
 CC is useful for treating or preventing a wide variety of disorders such as
 CC hemophilia, glycogen storage deficiency type 1A, Pompe deficiency,
 CC galactosmia, phenylketonuria, Crigler-Najjar disease, gout and Lesch-
 CC Nyan syndrome, Zellweger syndrome, ischemic diseases, atherosclerosis,
 CC thrombosis, embolisms, Parkinson's disease, congestive heart failure,
 CC cancer, inflammatory and immune disorders, muscular dystrophies, and
 CC diabetes. This is the amino acid sequence of adeno-associated virus 1
 CC (AAV1) capsid protein VP1.

XX
 XX
 SQ Sequence 736 AA;

Query Match 100.0%; Score 3989; DB 9; Length 736;

Best Local Similarity 100.0%; Pred. No. 5.2e-311; Mismatches 0; Gaps 0;

Db 1 MAADGYLPMLENDLSGIREWMDLKPAPKPKANQKODDGRGLVPGYKYLPPFNGLD 60
 1 MAADGYLPMLENDLSGIREWMDLKPAPKPKANQKODDGRGLVPGYKYLPPFNGLD 60
 QY 61 KGEFVNADAALAEHDAKAYDOQLKAGDNPYLRYNHADAEPFERQEDTSPFGNIGRAVFO 120
 61 KGEFVNADAALAEHDAKAYDOQLKAGDNPYLRYNHADAEPFERQEDTSPFGNIGRAVFO 120
 Db 61 KGEFVNADAALAEHDAKAYDOQLKAGDNPYLRYNHADAEPFERQEDTSPFGNIGRAVFO 120
 QY 121 AKKRVLEPLGLVEGAKTAPGKKRPVEQSPQEPDSSSGIGTGOQPAKKRLNFGQTDSE 180
 121 AKKRVLEPLGLVEGAKTAPGKKRPVEQSPQEPDSSSGIGTGOQPAKKRLNFGQTDSE 180
 Db 121 AKKRVLEPLGLVEGAKTAPGKKRPVEQSPQEPDSSSGIGTGOQPAKKRLNFGQTDSE 180
 QY 181 SVDPPOPLGEPBPATPAVGPPTTMASSGGAPMADNNEGADVGNSGMMHCSTYLGDRVI 240
 181 SVDPPOPLGEPBPATPAVGPPTTMASSGGAPMADNNEGADVGNSGMMHCSTYLGDRVI 240
 Db 181 SVDPPOPLGEPBPATPAVGPPTTMASSGGAPMADNNEGADVGNSGMMHCSTYLGDRVI 240
 QY 241 TTSRTALPTYNHLYKQIASSASTGASNDNHFGYSTPMGYPFNFPHCFHSPRDQRL 300
 241 TTSRTALPTYNHLYKQIASSASTGASNDNHFGYSTPMGYPFNFPHCFHSPRDQRL 300
 Db 241 TTSRTALPTYNHLYKQIASSASTGASNDNHFGYSTPMGYPFNFPHCFHSPRDQRL 300
 QY 301 INNNMGRPRKLNKLFNIOVKEVTNDGVTIANNLITVQVESDSEYQLPYLGSANQ 360
 301 INNNMGRPRKLNKLFNIOVKEVTNDGVTIANNLITVQVESDSEYQLPYLGSANQ 360
 Db 301 INNNMGRPRKLNKLFNIOVKEVTNDGVTIANNLITVQVESDSEYQLPYLGSANQ 360
 QY 361 GCLPFPADVPMTFOYGYLTLANSSQAVGRSSFCLEFPBQMRTGNFFSTFEFVP 420
 361 GCLPFPADVPMTFOYGYLTLANSSQAVGRSSFCLEFPBQMRTGNFFSTFEFVP 420
 Db 361 GCLPFPADVPMTFOYGYLTLANSSQAVGRSSFCLEFPBQMRTGNFFSTFEFVP 420
 QY 421 FHSSVYASQSLDRMLANPLIDQYLYLNRTOQSSAQNKDLLFRGSPAGMSVOPKMWLP 480
 421 FHSSVYASQSLDRMLANPLIDQYLYLNRTOQSSAQNKDLLFRGSPAGMSVOPKMWLP 480
 Db 421 FHSSVYASQSLDRMLANPLIDQYLYLNRTOQSSAQNKDLLFRGSPAGMSVOPKMWLP 480
 QY 481 GPCYRQGRVSKTKTDNNNSNFTWTGASKYNLNGRESIINPGTAMASHKDEKFPMSGV 540
 481 GPCYRQGRVSKTKTDNNNSNFTWTGASKYNLNGRESIINPGTAMASHKDEKFPMSGV 540
 Db 481 GPCYRQGRVSKTKTDNNNSNFTWTGASKYNLNGRESIINPGTAMASHKDEKFPMSGV 540
 QY 541 MIREKESAGASNTLADVMITDEEIKATNPVATERGTVAVNVOSSSTDATDVAAMG 600
 541 MIREKESAGASNTLADVMITDEEIKATNPVATERGTVAVNVOSSSTDATDVAAMG 600
 Db 541 MIREKESAGASNTLADVMITDEEIKATNPVATERGTVAVNVOSSSTDATDVAAMG 600
 QY 601 ALPQWVQDRDVLVQGPIMAKIPTHDGHFHSPPLMGFGGLKNPPQILIKTVPANPPA 660
 601 ALPQWVQDRDVLVQGPIMAKIPTHDGHFHSPPLMGFGGLKNPPQILIKTVPANPPA 660
 Db 601 ALPQWVQDRDVLVQGPIMAKIPTHDGHFHSPPLMGFGGLKNPPQILIKTVPANPPA 660

QY 661 EFSATKASPTITQYSTGVSEIEMELOKENSKRANBEVOYTSNYSKSNVDFTVNNGL 720
 661 EFSATKASPTITQYSTGVSEIEMELOKENSKRANBEVOYTSNYSKSNVDFTVNNGL 720
 Db 661 EFSATKASPTITQYSTGVSEIEMELOKENSKRANBEVOYTSNYSKSNVDFTVNNGL 720
 QY 721 YTEPRPIGRITRYLTPRL 736
 721 YTEPRPIGRITRYLTPRL 736
 Db 721 YTEPRPIGRITRYLTPRL 736

RESULT 6

AD227069
 ID AD227069 standard; protein; 736 AA.

AC AD227069;

DT 30-JUN-2005 (first entry)

XX Adeno-associated virus protein SEQ ID NO 219.

XX rheumatoid arthritis; multiple sclerosis; sarcoidosis; diabetes;

XX scleroderma; psoriasis; vasculitis; Crohn's disease; hemophilia; HIV;

XX bacterial infection; cancer; ulcerative colitis; antirheumatic;

XX antiafatic; neuroprotective; antiinflammatory; antidiabetic;

XX antipariatic; vasotropic; gastrointestinal-gen; hemostatic; anti-HIV;

XX virucide; antibacterial; cytostatic; antitumor; dermatological.

XX Adeno-associated virus.

XX WO200503321-A2.

XX 30-SEP-2004; 2004WO-US028817.

XX 30-SEP-2003; 2003US-0508226P.

XX 29-APR-2004; 2004US-0566546P.

XX (VYPE-) UNIV PENNSYLVANIA.

XX Wilson JM, Gao G, Alvira MR, Vandenbergh LH;

XX WPI; 2005-285437/29.

XX New adeno-associated virus (AAV) clade comprising at least three AAV

XX members, useful for preventing and/or treating arthritis, multiple

XX sclerosis, diabetes, scleroderma, psoriasis, hemophilia, HIV, bacterial

XX infection and cancer.

XX Disclosure; SEQ ID NO 219; 569pp; English.

XX The invention relates to an adeno-associated virus (AAV) clade comprising

XX at least three AAV members, where each member of the AAV clade is

XX phylogenetically related as determined using a Neighbor-joining heuristic

XX by a bootstrap value of at least 75 % per 1000 isolates and a Poisson

XX correction distance measurement of no more than 0.05. The methods and

XX compositions of the present invention are useful for the prevention

XX and/or treatment of rheumatoid arthritis, multiple sclerosis,

XX sarcoidosis, diabetes, scleroderma, psoriasis, vasculitis, Crohn's

XX disease, hemophilia, HIV, bacterial infection, cancer and ulcerative

```

Db      |||||
61      KGEPVNAADAAALEHDKAYDQOLKAGNPLRLYNHDAEFQERLQEDTSTFGNLGRAVQ 120
Qy      121 AKKRVLEPLGLVEGAKTAPGKKRPVEQSPQEPDSSSGIGKTQOQPAKKRLNFGQTGDE 180
Db      121 AKKRVLEPLGLVEGAKTAPGKKRPVEQSPQEPDSSSGIGKTQOQPAKKRLNFGQTGDE 180
Qy      181 SVDPQPLGEPPTPAVGPPTMASGGGAPMADNNEGADVGNASGNWCHDSTWLGDRVI 240
Db      181 SVDPQPLGEPPTPAVGPPTMASGGGAPMADNNEGADVGNASGNWCHDSTWLGDRVI 240
Qy      241 TTSTRTWALPTYNHLYKOISSASTGASNDNHFGYSTPMGYDFPNRFGHFSRDMQRL 300
Db      241 TTSTRTWALPTYNHLYKOISSASTGASNDNHFGYSTPMGYDFPNRFGHFSRDMQRL 300
Qy      301 INNMGFRPKRLNFKLFNIQVKEVTNDGVTIANNLSTVQVFSDEYOLPYVLGSAHQ 360
Db      301 INNMGFRPKRLNFKLFNIQVKEVTNDGVTIANNLSTVQVFSDEYOLPYVLGSAHQ 360
Qy      361 GCLPPPADVFMIPQGYLTLLNNGSQAVGSSFCLEFPSSQMLRTGNNTFSTYFEVP 420
Db      361 GCLPPPADVFMIPQGYLTLLNNGSQAVGSSFCLEFPSSQMLRTGNNTFSTYFEVP 420
Qy      421 FHSSYAHOSGLDRLMPLIDQYLYLNRTQNGSAGNKKLLFRSGSPAGMSVQPKWLP 480
Db      421 FHSSYAHOSGLDRLMPLIDQYLYLNRTQNGSAGNKKLLFRSGSPAGMSVQPKWLP 480
Qy      481 GPCYRQORVSKTIDNNNSNFTWTGASKYMLNGRESIINPGTAMASHKDEDEKFFPMGV 540
Db      481 GPCYRQORVSKTIDNNNSNFTWTGASKYMLNGRESIINPGTAMASHKDEDEKFFPMGV 540
Qy      541 MIFGESAGASNTALDNVMTTDEBEIKATNPVATERGTVAVNFQSSSTDPATGDVHAMG 600
Db      541 MIFGESAGASNTALDNVMTTDEBEIKATNPVATERGTVAVNFQSSSTDPATGDVHAMG 600
Qy      601 ALGGMWODRDVYLOGIWKIPHTDGHFHSPLMGFGLKNPPQLIKNTVPAPAPPA 660
Db      601 ALGGMWODRDVYLOGIWKIPHTDGHFHSPLMGFGLKNPPQLIKNTVPAPAPPA 660
Qy      661 EFSATKASPTITQSTQVSEIEWELQKENSCKRMNBEVOYTSNVAKSANVDFVDNGL 720
Db      661 EFSATKASPTITQSTQVSEIEWELQKENSCKRMNBEVOYTSNVAKSANVDFVDNGL 720
Qy      721 YTEPRPIGTIRYLRPL 736
Db      721 YTEPRPIGTIRYLRPL 736

```

RESULT 7

```

AD227012
ID      AD227012 standard; protein; 736 AA.
XX
AC      AD227012;
XX
DT      30-JUN-2005 (first entry)
XX
DE      Adeno-associated virus protein SEQ ID NO 162.
XX
KW      rheumatoid arthritis; multiple sclerosis; sarcoidosis; diabetes;
KW      scleroderma; psoriasis; vasculitis; Crohn's disease; hemophilia; HIV;
KW      bacterial infection; cancer; ulcerative colitis; antirheumatic;
KW      anticholinergic; neuroprotective; antiinflammatory; antidiabetic;
KW      antipsoriatic; vasculitic; gastrointestinal-gen.; hemostatic; anti-HIV;
KW      virucide; antibacterial; cytostatic; antitumor; dermatological.
XX
OS      Adeno-associated virus.
XX
PN      WO2005033321-A2.
XX
PD      14-APR-2005.
XX
PF      30-SEP-2004; 2004WO-US028817.
XX

```

```

PR      30-SEP-2003; 2003US-0508226P.
PR      29-APR-2004; 2004US-0566546P.
XX
XX      (type-) UNIV PENNSYLVANIA.
PI      Wilson JM, Gao G, Alvira MR, Vandenbergh LH;
XX      WPI; 2005-285437/29.
XX
XX      New adeno-associated virus (AAV) clade comprising at least three AAV
XX      members, useful for preventing and/or treating arthritis, multiple
XX      sclerosis, diabetes, scleroderma, psoriasis, hemophilia, HIV, bacterial
XX      infection and cancer.
PS      Disclosure, SEQ ID NO 162; 569pp; English.
XX
XX      The invention relates to an adeno-associated virus (AAV) clade comprising
XX      at least three AAV members, where each member of the AAV clade is
XX      phylogenetically related as determined using a Neighbor-joining heuristic
XX      by a bootstrap value of at least 75 % per 1000 isolates and a Poisson
XX      correction distance measurement of no more than 0.05. The methods and
XX      compositions of the present invention are useful for the prevention
XX      and/or treatment of rheumatoid arthritis, multiple sclerosis,
XX      Crohn's disease, diabetes, scleroderma, psoriasis, vasculitis, Crohn's
XX      disease, hemophilia, HIV, bacterial infection, cancer and ulcerative
XX      colitis. The present sequence represents the amino acid sequence of an
XX      adeno-associated virus protein.
XX
XX      Sequence 736 AA:
XX
XX      Query Match          100.0%; Score 3989; DB 9; Length 736;
XX      Best Local Similarity 100.0%; Pred No. 5.2e-311;
XX      Matches 736; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY 601 ALPGMWQDRDVLVLOGPIWAKIPHTDGHFHPSPIMGFGFLKNPPQILLIKNTVPANPPA 660
DB 601 ALPGMWQDRDVLVLOGPIWAKIPHTDGHFHPSPIMGFGFLKNPPQILLIKNTVPANPPA 660
QY 661 EFSATKRFASFTITQSTGQVSEIEMELQKNSKRMNEVOYTSNYAKSANVDFTVDNGL 720
DB 661 EFSATKRFASFTITQSTGQVSEIEMELQKNSKRMNEVOYTSNYAKSANVDFTVDNGL 720
QY 721 YTEPRPIGTRYLTRPL 736
DB 721 YTEPRPIGTRYLTRPL 736
RESULT 8
ADV67506
ID ADV67506 standard; protein: 735 AA.
XX
AC ADV67506;
XX
DT 10-MAR-2005 (first entry)
XX
DE Amino acid sequence of the capsid protein of AAV serotype 1.
XX
KW antitripteriosclerotic; antilipemic; gene therapy; cholesterol;
KW apolipoprotein E; apoe; apolipoprotein A; apoa; atherosclerosis;
KW lipoprotein defect; capsid protein.
XX
OS Adeno-associated virus.
XX
PN WO2004108922-A2.
XX
PD 16-DEC-2004.
XX
PF 23-APR-2004; 2004WO-US010965.
XX
PR 25-APR-2003; 2003JUS-0465293P.
XX
PA (UYPE-) UNIV PENNSYLVANIA.
XX
PI Rader DJ, Wilson JM;
XX
DR WPI; 2005-031700/03.
XX
PT Lowering total cholesterol levels and treating atherosclerosis in a
PT subject comprises delivering a recombinant adeno-associated virus (AAV)
PT comprising an AAV serotype capsid protein or a gene encoding human
PT apolipoprotein E (apoE) or apoA.
XX
PS Disclosure; SEQ ID NO 4; 699p; English.
XX
CC The specification describes a method for lowering total cholesterol
CC levels in a subject. The method comprises delivering to the subject a
CC recombinant adeno-associated virus (AAV) comprising a gene encoding a
CC human apolipoprotein E (apoE) or apoA under the control of a regulatory
CC control sequence which directs expression of the gene. The recombinant
CC AAV also comprises a capsid protein selected from an AAV serotype (e.g.,
CC AAV7 or AAV8) which preferentially expresses high levels of transgene in
CC liver. A therapeutically effective amount of apoE or apoA expression is
CC obtained upon delivery of low dose of AAV. The method of the invention is
CC useful for lowering total cholesterol levels in a subject, e.g. for
CC treating atherosclerosis, and for correcting defects in lipoprotein. The
CC present sequence represents a capsid protein of AAV serotype 1 (AAV1),
CC which may be used in recombinant AAV vectors of the invention.
XX
SQ Sequence 735 AA.
Query Match 99.9%; Score 3985; DB 9; Length 735;
Best Local Similarity 100.0%; Pred. No. 1.1e-310;
Matches 735; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MAADGYLPDLIEDNLSEIGREWMDLKGAARKPKANQOKODDGRGLVLPGYKYLGFPMGLD 60
|||||

DB 1 MAADGYLPDLIEDNLSEIGREWMDLKGAARKPKANQOKODDGRGLVLPGYKYLGFPMGLD 60
QY 61 KGEFVNADAALAEHDKAYDOOLKAGDNPYLRYNHADAEEFERLOEDTSPFGNIGRAVFQ 120
DB 61 KGEFVNADAALAEHDKAYDOOLKAGDNPYLRYNHADAEEFERLOEDTSPFGNIGRAVFQ 120
QY 121 AKKRVLEPLGVEGAKTAPGKKRPVEQSPQEPSPSSSGIGTGQOPAKKRLNFCQTGSE 180
DB 121 AKKRVLEPLGVEGAKTAPGKKRPVEQSPQEPSPSSSGIGTGQOPAKKRLNFCQTGSE 180
QY 181 SVDPPOPGLGEPPTPAVGPPTMASGGGAPMADNNEGADGVGNAGMWHCSTLIGDRV 240
DB 181 SVDPPOPGLGEPPTPAVGPPTMASGGGAPMADNNEGADGVGNAGMWHCSTLIGDRV 240
QY 241 TTSRTWALPTYNHLYKQISASTGASNDNHYFGYSTPMGYFPDFNRFHCFSPRDQRL 300
DB 241 TTSRTWALPTYNHLYKQISASTGASNDNHYFGYSTPMGYFPDFNRFHCFSPRDQRL 300
QY 301 INNNMGFRPKRLNFKLFNIQVKEYTNDGYTTIANNTSTYQVSDSEYQLPYLGSAAHQ 360
DB 301 INNNMGFRPKRLNFKLFNIQVKEYTNDGYTTIANNTSTYQVSDSEYQLPYLGSAAHQ 360
QY 361 GCLPEPPADVPMIPOGYLTLNNGSOAVGRSSPYCLEYPPSQMLRTGNNFTFSYTFEEVP 420
DB 361 GCLPEPPADVPMIPOGYLTLNNGSOAVGRSSPYCLEYPPSQMLRTGNNFTFSYTFEEVP 420
QY 421 FHSYSYASQSLDRIMNPLIDQYLYLNRTONQSSAQNKDLLFSRGSFAGMSVOPKMWLP 480
DB 421 FHSYSYASQSLDRIMNPLIDQYLYLNRTONQSSAQNKDLLFSRGSFAGMSVOPKMWLP 480
QY 481 GPCYRQGRVSKTKTDNNNSNFTWTGASKYLNNGRESIINPGTAMASHKDEDEKFFPMGCV 540
DB 481 GPCYRQGRVSKTKTDNNNSNFTWTGASKYLNNGRESIINPGTAMASHKDEDEKFFPMGCV 540
QY 541 MIFKESAGASNTALDVMITDDEEIKATNPVATERGTYAVNFOSSSTDPAIDVAMG 600
DB 541 MIFKESAGASNTALDVMITDDEEIKATNPVATERGTYAVNFOSSSTDPAIDVAMG 600
QY 601 ALPGMWQDRDVLVLOGPIWAKIPHTDGHFHPSPIMGFGFLKNPPQILLIKNTVPANPPA 660
DB 601 ALPGMWQDRDVLVLOGPIWAKIPHTDGHFHPSPIMGFGFLKNPPQILLIKNTVPANPPA 660
QY 661 EFSATKRFASFTITQSTGQVSEIEMELQKNSKRMNEVOYTSNYAKSANVDFTVDNGL 720
DB 661 EFSATKRFASFTITQSTGQVSEIEMELQKNSKRMNEVOYTSNYAKSANVDFTVDNGL 720
QY 721 YTEPRPIGTRYLTRP 735
DB 721 YTEPRPIGTRYLTRP 735
RESULT 9
AD227007
ID AD227007 standard; protein: 736 AA.
XX
AC AD227007;
XX
DT 30-JUN-2005 (first entry)
XX
DE Adeno-associated virus protein SEQ ID NO 157.
XX
KW rheumatoid arthritis; multiple sclerosis; sarcoidosis; diabetes;
KW scleroderma; psoriasis; vasculitis; Crohn's disease; hemophilia; HIV;
KW bacterial infection; cancer; ulcerative colitis; antirheumatic;
KW antiarthritic; neuroprotective; antiinflammatory; antiadipatic;
KW antidiabetic; vasotrophic; gastrointestinal-gen.; hemostatic; anti-HIV;
KW virucide; antibacterial; cytostatic; antidiacer; dermatological.
XX
OS Adeno-associated virus.
XX
PN WO2005033321-A2.
XX
PD 14-APR-2005.

```
XX 30-SEP-2004; 2004MO-US028817.
PF 30-SEP-2003; 2003US-0508226P.
XX 29-APR-2004; 2004US-0566546P.
XX (UYPE-) UNIV PENNSYLVANIA.
XX Wilson JM, Gao G, Alvira KR, Vandenbergh LH;
XX WPI; 2005-285437/29.
XX
XX New adeno-associated virus (AAV) clade comprising at least three AAV
XX members, useful for preventing and/or treating arthritis, multiple
XX sclerosis, diabetes, scleroderma, psoriasis, hemophilia, HIV, bacterial
XX infection and cancer.
XX
XX Claim 19; SEQ ID NO 157; 569pp; English.
XX
XX The invention relates to an adeno-associated virus (AAV) clade comprising
XX at least three AAV members, where each member of the AAV clade is
XX phylogenetically related as determined using a Neighbor-Joining heuristic
XX by a bootstrap value of at least 75 % per 1000 isolates and a Poisson
XX correction distance measurement of no more than 0.05. The methods and
XX compositions of the present invention are useful for the prevention
XX and/or treatment of rheumatoid arthritis, multiple sclerosis,
XX sarcoidosis, diabetes, scleroderma, psoriasis, vasculitis, Crohn's
XX disease, hemophilia, HIV, bacterial infection, cancer and ulcerative
XX colitis. The present sequence represents the amino acid sequence of an
XX adeno-associated virus protein.
XX
XX Sequence 736 AA:
XX
Query Match          99.3%; Score 3968; DB 9; Length 736;
Best Local Similarity 99.3%; Pred. No. 2.5e-309;
Matches 731; Conservative 3; Mismatches 2; Indels 0; Gaps 0;
XX
XX 1 MAADGYLPDWLENDLSEGIREFWMDLKGAPKPRANQOKODGGLVLPGRYKLGPFNGLD 60
XX 1 MAADGYLPDWLENDLSEGIREFWMDLKGAPKPRANQOKODGGLVLPGRYKLGPFNGLD 60
XX
XX KGEPVNAADAAAEHDKAYDQOLKAGNPLYRNVHADEFQERLOEDTSGNLGRAVFO 120
XX 61 KGEPVNAADAAAEHDKAYDQOLKAGNPLYRNVHADEFQERLOEDTSGNLGRAVFO 120
XX
XX 121 AKKRVLEPLGLVEGAKTAPGKRPVQSPQEPDSSSGIGTQGPAAKRLNFGQTDSE 180
XX 121 AKKRVLEPLGLVEGAKTAPGKRPVQSPQEPDSSSGIGTQGPAAKRLNFGQTDSE 180
XX
XX 181 SVDPDQPLGPRPATPAVGTPTMASGGGAPMADNNEGADGVGNAGWHDSTWLGDRVI 240
XX 181 SVDPDQPLGPRPATPAVGTPTMASGGGAPMADNNEGADGVGNAGWHDSTWLGDRVI 240
XX
XX 181 SVDPDQPLGPRPATPAVGTPTMASGGGAPMADNNEGADGVGNAGWHDSTWLGDRVI 240
XX 181 SVDPDQPLGPRPATPAVGTPTMASGGGAPMADNNEGADGVGNAGWHDSTWLGDRVI 240
XX
XX 241 TTSRTTAAALPTYNHNLKQISSASTGASNDNHVGYSTPMGYDPFNRFCHFSRPRDQRL 300
XX 241 TTSRTTAAALPTYNHNLKQISSASTGASNDNHVGYSTPMGYDPFNRFCHFSRPRDQRL 300
XX
XX 301 INNMGFRPRRLNFKLFNIOVEVTTNDGVTTIANNLSTVQVSEDEYQLPYVLSAQ 360
XX 301 INNMGFRPRRLNFKLFNIOVEVTTNDGVTTIANNLSTVQVSEDEYQLPYVLSAQ 360
XX
XX 361 GCLPPPPADVFMIPOGYTLTLNNGSQAVGRSSFCLCYFPSONMLRTNNFTFSTFEVP 420
XX 361 GCLPPPPADVFMIPOGYTLTLNNGSQAVGRSSFCLCYFPSONMLRTNNFTFSTFEVP 420
XX
XX 421 FHSSYAHOSQSLDRMLNPLIDQYLYLNRTONOSGSAONKLLFSGSPAGMSVQPKWMLP 480
XX 421 FHSSYAHOSQSLDRMLNPLIDQYLYLNRTONOSGSAONKLLFSGSPAGMSVQPKWMLP 480
XX
XX 481 GPCYRQORVSKTKTDNNNSNFTWTGASKYNLNGRESIINPGTAMASHKODEKFFPMG 540
XX 481 GPCYRQORVSKTKTDNNNSNFTWTGASKYNLNGRESIINPGTAMASHKODEKFFPMG 540
XX
XX 481 GPCYRQORVSKTKTDNNNSNFTWTGASKYNLNGRESIINPGTAMASHKODEKFFPMG 540
XX 481 GPCYRQORVSKTKTDNNNSNFTWTGASKYNLNGRESIINPGTAMASHKODEKFFPMG 540
```

```
QY 541 MIFGKESAGASNTALNNMTDEEIKATNPVATERFGVAVVFOSSSTDPAQDVHAMG 600
DB 541 MIFGKESAGASNTALNNMTDEEIKATNPVATERFGVAVVFOSSSTDPAQDVHAMG 600
QY 601 ALPGMWQODRDVYLQGPIMAKIPHTDGHFPPSPLMGFGGLKNPPQILIKNTVPANPPA 660
DB 601 ALPGMWQODRDVYLQGPIMAKIPHTDGHFPPSPLMGFGGLKNPPQILIKNTVPANPPA 660
QY 661 EFSATKPFASFTIQYSTGVSVSEIEMELOKENSRRKNPEVOYTSNVAKSANVDFTVNGGL 720
DB 661 EFSATKPFASFTIQYSTGVSVSEIEMELOKENSRRKNPEVOYTSNVAKSANVDFTVNGGL 720
QY 721 YTEPRPIGTRYLRLPL 736
DB 721 YTEPRPIGTRYLRLPL 736
XX
XX RESULT 10
XX AAB59847
XX ID AAB59847 standard; protein; 736 AA.
XX
XX AAB59847;
XX AC 28-MAR-2001 (first entry)
XX
XX DE AAV6 capsid protein VP1.
XX
XX KM AAV6; gene therapy; AAV viral vector; cystic fibrosis; cancer; AIDS;
XX KM atherosclerosis; sickle cell anaemia; thalassemia;
XX KM blood clotting disorder; diabetes; capsid protein VP1.
XX
XX OS Adeno associated virus.
XX
XX PN US6156303-A.
XX
XX PD 05-DEC-2000.
XX
XX PF 11-JUN-1997; 97US-00873168.
XX
XX PR 11-JUN-1997; 97US-00873168.
XX
XX PA (UNIW ) UNIV WASHINGTON.
XX
XX PI Russell DW, Rutledge EA;
XX
XX DR WPI; 2001-060164/07.
XX
XX PT Adeno-associated virus serotype 6 and viral vector derived from it for
XX PT gene therapy of cystic fibrosis, cancer, acquired immunodeficiency
XX PT syndrome, sickle cell anemia, thalassemia and diabetes.
XX
XX PS Disclosure; Fig 2; 50pp; English.
XX
XX The present invention relates to adeno-associated virus serotypes. The
XX present sequence is capsid protein VP1 of one such serotype (AAV6). AAV6
XX can be used to construct AAV viral vectors for use in gene therapy for a
XX range of disorders: cystic fibrosis, cancer, AIDS, atherosclerosis,
XX sickle cell anaemia, thalassemia, blood clotting disorders and diabetes.
XX The AAV viral vectors have increased transduction efficiency of a
XX particular host cell as the AAV virion containing the AAV vector genome
XX can be modified to express a capsid protein of an AAV serotype that
XX transduces the selected host cell
XX
XX Sequence 736 AA:
XX
Query Match          99.3%; Score 3963; DB 4; Length 736;
Best Local Similarity 99.2%; Pred. No. 6.4e-309;
Matches 730; Conservative 3; Mismatches 3; Indels 0; Gaps 0;
XX
XX 1 MAADGYLPDWLENDLSEGIREFWMDLKGAPKPRANQOKODGGLVLPGRYKLGPFNGLD 60
XX 1 MAADGYLPDWLENDLSEGIREFWMDLKGAPKPRANQOKODGGLVLPGRYKLGPFNGLD 60
```



```

Qy 61 KGEVNAADAALAEHDKAYDOOLKAGDNPLYLRVNHADAERFQERLOEDTSFGGNGRAVFO 120
Db 61 KGEVNAADAALAEHDKAYDOOLKAGDNPLYLRVNHADAERFQERLOEDTSFGGNGRAVFO 120
Qy 121 AKKRVLEPGLVEGAKTAPGKKRPVEQSPOEPDSSSGIGKTGOQPAKKRLNFCQTGDS 180
Db 121 AKKRVLEPGLVEGAKTAPGKKRPVEQSPOEPDSSSGIGKTGOQPAKKRLNFCQTGDS 180
Qy 181 SVDPQPLGEBPATPAAVGPTTMASSGGAPMADNNEGADGVGNASGMWHCSTWLGDRVI 240
Db 181 SVDPQPLGEBPATPAAVGPTTMASSGGAPMADNNEGADGVGNASGMWHCSTWLGDRVI 240
Qy 241 TTSRTWALPTYNNHLKYQISSASTGASNDNHFGYSTPMGYFPFNNFCHFSRDMQRL 300
Db 241 TTSRTWALPTYNNHLKYQISSASTGASNDNHFGYSTPMGYFPFNNFCHFSRDMQRL 300
Qy 301 INNMGRPRKLNFKLFNIOVKEVTTNDGVTIANNLSTYQVPSDSEYQLPYVLSAHQ 360
Db 301 INNMGRPRKLNFKLFNIOVKEVTTNDGVTIANNLSTYQVPSDSEYQLPYVLSAHQ 360
Qy 361 GCLPPPADVFMIPQYGLTLNNGSOAVGRSSFYCLEYFSPQMLRTGNNFTFSYTFEEVP 420
Db 361 GCLPPPADVFMIPQYGLTLNNGSOAVGRSSFYCLEYFSPQMLRTGNNFTFSYTFEEVP 420
Qy 421 FHSSTYAHQSGLDRLMNPLIDQYLYLNRTQNGSSAQNKDLFSGSPAGMSVOPKMWLP 480
Db 421 FHSSTYAHQSGLDRLMNPLIDQYLYLNRTQNGSSAQNKDLFSGSPAGMSVOPKMWLP 480
Qy 481 GPCYRQORVSKTKTDNNNSNFTWTGASKYNLNGRESIINPGTAMASHKDDKDFPMSGV 540
Db 481 GPCYRQORVSKTKTDNNNSNFTWTGASKYNLNGRESIINPGTAMASHKDDKDFPMSGV 540
Qy 541 MIFGESAGASNTALDNVMTDEBEIKATNPVATERFGYAVNFQSSSTDPATGDVHAMG 600
Db 541 MIFGESAGASNTALDNVMTDEBEIKATNPVATERFGYAVNFQSSSTDPATGDVHAMG 600
Qy 601 ALRGMTWQDRDVIYQGIWAKIPHTDGHFHSPLMGSGGLKNPPQILIKTTPVNPANPPA 660
Db 601 ALRGMTWQDRDVIYQGIWAKIPHTDGHFHSPLMGSGGLKNPPQILIKTTPVNPANPPA 660
Qy 661 EFSATKASFTIOTSGOVSEIEMELOKENSKRMNPEVOYTSNVAASANDPFTVDNGL 720
Db 661 EFSATKASFTIOTSGOVSEIEMELOKENSKRMNPEVOYTSNVAASANDPFTVDNGL 720
Qy 721 YTEPRPIGTRYLTRPL 736
Db 721 YTEPRPIGTRYLTRPL 736

```

```

RESULT 11
ADE76566
ID ADE76566 standard; protein; 736 AA.
XX
AC ADE76566;
XX
DT 29-JAN-2004 (first entry)
XX
DE Adeno-associated virus (AAV) related protein, SEQ ID NO 65.
XX
KW adeno-associated virus; AAV; cytosolic; antipsoaritic; antirheumatic;
KW antidiabetic; neuroprotective; antidiabetic; antithyroid;
KW dermatologic; antiinflammatory; gene therapy; vaccine;
KW hyperproliferative; cancer; psoriasis; autoimmune disease;
KW rheumatoid arthritis; multiple sclerosis; diabetes;
KW autoimmune thyroiditis; scleroderma; Crohn's disease.
XX
OS unidentified.
XX
PN EP1310571-A2.
XX
PD 14-MAY-2003.
XX
PF 12-NOV-2002; 2002EP-00257826.

```

```

XX 13-NOV-2001; 2001US-0350607P.
PR 17-DEC-2001; 2001US-0341117P.
PR 01-MAY-2002; 2002US-0377066P.
PR 05-JUN-2002; 2002US-0386757P.
XX
PA (UNP-) UNIV PENNSYLVANIA.
XX
PI Gao G, Wilson JM, Alvira M;
XX
DR WPI; 2003-450984/43.
XX
PT Detecting adeno-associated virus sequences in a sample, useful for e.g.
PT preventing or treating hyperproliferative or autoimmune diseases,
PT comprising subjecting a sample having a DNA to amplification via
PT polymerase chain reaction.
XX
PS Disclosure; SEQ ID NO 65; 419pp; English.
XX
CC The invention relates to a novel method for detecting adeno-associated
CC virus (AAV) sequences in a sample, which comprises subjecting a sample
CC containing a DNA to amplification via a polymerase chain reaction (PCR).
CC The AAV sequence have the following activities: cytostatic,
CC antipsoaritic, antirheumatic, antidiabetic, neuroprotective,
CC antidiabetic, antithyroid, dermatologic, and antiinflammatory. The AAV
CC sequence can be used in gene therapy or as part of a vaccine to treat
CC disorders. The method is useful in detecting and/or identifying AAV
CC sequences and isolating novel sequences that are identified. The
CC sequences may be used e.g. for preventing or treating hyperproliferative
CC conditions such as cancers and psoriasis, and other autoimmune diseases
CC like rheumatoid arthritis, multiple sclerosis, diabetes, autoimmune
CC thyroiditis, scleroderma or Crohn's disease. This sequence represents an
CC AAV related protein sequence of the invention.
XX
SQ Sequence 736 AA:

```

```

Query Match 99.3%; Score 3963; DB 7; Length 736;
Best Local Similarity 99.2%; Pred. No. 6.4e-309;
Matches 730; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

```

```

Qy 1 MAADGYPLDMLENDLSRGIREWMDLKPAPKPNQKODDGRGLVPGYKYLGPENLD 60
Db 1 MAADGYPLDMLENDLSRGIREWMDLKPAPKPNQKODDGRGLVPGYKYLGPENLD 60
Qy 61 KGEVNAADAALAEHDKAYDOOLKAGDNPLYLRVNHADAERFQERLOEDTSFGGNGRAVFO 120
Db 61 KGEVNAADAALAEHDKAYDOOLKAGDNPLYLRVNHADAERFQERLOEDTSFGGNGRAVFO 120
Qy 121 AKKRVLEPGLVEGAKTAPGKKRPVEQSPOEPDSSSGIGKTGOQPAKKRLNFCQTGDS 180
Db 121 AKKRVLEPGLVEGAKTAPGKKRPVEQSPOEPDSSSGIGKTGOQPAKKRLNFCQTGDS 180
Qy 181 SVDPQPLGEBPATPAAVGPTTMASSGGAPMADNNEGADGVGNASGMWHCSTWLGDRVI 240
Db 181 SVDPQPLGEBPATPAAVGPTTMASSGGAPMADNNEGADGVGNASGMWHCSTWLGDRVI 240
Qy 241 TTSRTWALPTYNNHLKYQISSASTGASNDNHFGYSTPMGYFPFNNFCHFSRDMQRL 300
Db 241 TTSRTWALPTYNNHLKYQISSASTGASNDNHFGYSTPMGYFPFNNFCHFSRDMQRL 300
Qy 301 INNMGRPRKLNFKLFNIOVKEVTTNDGVTIANNLSTYQVPSDSEYQLPYVLSAHQ 360
Db 301 INNMGRPRKLNFKLFNIOVKEVTTNDGVTIANNLSTYQVPSDSEYQLPYVLSAHQ 360
Qy 361 GCLPPPADVFMIPQYGLTLNNGSOAVGRSSFYCLEYFSPQMLRTGNNFTFSYTFEEVP 420
Db 361 GCLPPPADVFMIPQYGLTLNNGSOAVGRSSFYCLEYFSPQMLRTGNNFTFSYTFEEVP 420
Qy 421 FHSSTYAHQSGLDRLMNPLIDQYLYLNRTQNGSSAQNKDLFSGSPAGMSVOPKMWLP 480
Db 421 FHSSTYAHQSGLDRLMNPLIDQYLYLNRTQNGSSAQNKDLFSGSPAGMSVOPKMWLP 480
Qy 481 GPCYRQORVSKTKTDNNNSNFTWTGASKYNLNGRESIINPGTAMASHKDDKDFPMSGV 540

```

```

DB      481 GPCYRQORVSKTTDNNNSFTWTGASKYVLRRESIINGGTAMASHKDDKDFPMSGV 540
QY      541 MIFGKSAGASNTALDNVMTTDEEBEIKATNPVATERFGTAAVNFQSSSTDPATGDVHAMG 600
DB      541 MIFGKSAGASNTALDNVMTTDEEBEIKATNPVATERFGTAAVNFQSSSTDPATGDVHAMG 600
QY      601 ALPGMWQDRDVLQGPIMAKIPHTDGHFSPPLMGFGFKHPPLQILKNTPEVPANPPA 660
DB      601 ALPGMWQDRDVLQGPIMAKIPHTDGHFSPPLMGFGFKHPPLQILKNTPEVPANPPA 660
QY      661 EFSATKFASTITQYSTGQVSEIEMELQKNSKRMPNEVOYTSNYAKSANVDFTVNNGL 720
DB      661 EFSATKFASTITQYSTGQVSEIEMELQKNSKRMPNEVOYTSNYAKSANVDFTVNNGL 720
QY      721 YTEPRPIGTRYLTRPL 736
DB      721 YTEPRPIGTRYLTRPL 736

RESULT 12
ADV70293
ID      ADV70293 standard; protein; 736 AA.
XX
AC      ADV70293;
XX
DT      10-MAR-2005 (first entry)
XX
DE      Primate adeno-associated virus 6 capsid protein VP1.
XX
KM      immunosuppressive; gene therapy; immunogenicity; virus inactivation;
KM      hemophilia; Pepck deficiency; galactosemia; phenylketonuria;
KM      Crigler-Najjar syndrome; Gout; Zellweger syndrome; ischemia;
KM      atherosclerosis; thrombosis; embolism; Parkinson's disease;
KM      congestive heart failure; cancer; inflammation; immune disorder;
KM      muscular dystrophy; diabetes; VPI.
XX
OS      Adeno-associated virus 6.
XX
PN      WO2004112727-A2.
XX
PD      29-DEC-2004.
XX
PF      21-JUN-2004; 2004MO-US019884.
PR      19-JUN-2003; 2003JUS-0480395P.
PR      30-APR-2004; 2004US-0567310P.
PR      03-JUN-2004; 2004US-0576501P.
XX
PA      (AVIG-) AVIGEN INC.
PI      Arbetman AE, Colosi P, Lochrie MA, Surosky RT;
XX
DR      WPI; 2005-048755/05.
XX
PT      New mutated adeno-associated virus (AAV) capsid protein that when present
PT      in an AAV virion imparts decreased immunoreactivity to the virion as
PT      compared to the corresponding wild-type virion, useful for treating e.g.
PT      hemophilia.
XX
PS      Example 5; SEQ ID NO 19; 136pp; English.
XX
CC      The invention describes a mutated adeno-associated virus (AAV) capsid
CC      protein that when present in an AAV virion imparts decreased
CC      immunoreactivity to the virion as compared to the corresponding wild-type
CC      virion. Also described are: a polynucleotide encoding the mutated protein
CC      above; a recombinant AAV virion comprising the mutated protein above; and
CC      delivering a recombinant AAV virion to a cell or tissue of a vertebrate
CC      subject. The recombinant AAV virion is useful for delivering a
CC      heterologous nucleic acid molecule to a cell or tissue of a vertebrate
CC      subject, where the protein encoded by the heterologous nucleic acid
CC      molecule is expressed at a level that provides a therapeutic effect,
CC      where the recombinant AAV virion may comprise a non-primate, mammalian

```

```

CC      AAV capsid protein that when present in an AAV virion imparts decreased
CC      immunoreactivity to the virion as compared to immunoreactivity of primate
CC      AAV-2, and the heterologous nucleic acid molecule, where the heterologous
CC      nucleic acid molecule encodes a therapeutic protein and is operably
CC      linked to control elements capable of directing the in vivo transcription
CC      and translation of the protein. The protein or the recombinant AAV virion
CC      is useful for treating or preventing a wide variety of disorders such as
CC      hemophilia, glycogen storage deficiency type 1a, Pepck deficiency,
CC      galactosemia, phenylketonuria, Crigler-Najjar disease, Gout and Lesch-
CC      Nyan syndrome, Zellweger syndrome, ischemic diseases, atherosclerosis,
CC      thrombosis, embolisms, Parkinson's disease, congestive heart failure,
CC      cancer, inflammatory and immune disorders, muscular dystrophies, and
CC      diabetes. This is the amino acid sequence of adeno-associated virus 6
CC      (AAV6) capsid protein VP1.
XX
SQ      Sequence 736 AA;
XX
Query Match          99.3%; Score 3963; DB 9; Length 736;
Best Local Similarity 99.2%; Pred. No. 6,4e-309;
Matches 730; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY      1 MAADGYLPWLIEDNLSEGIREFWMDLKPAPKPAKNOQKDDGGLVLPQYKYLGPFGND 60
DB      1 MAADGYLPWLIEDNLSEGIREFWMDLKPAPKPAKNOQKDDGGLVLPQYKYLGPFGND 60
QY      61 KGPVNAADAALAEHDKAVDQOLKAGDNPYLRYNHADAEPQERLOEDTSFGNLAGRAVFQ 120
DB      61 KGPVNAADAALAEHDKAVDQOLKAGDNPYLRYNHADAEPQERLOEDTSFGNLAGRAVFQ 120
QY      121 AKKRVLEPLGLVEGAKTAPGKKRPVEQSPQEBSSSGIOKTDGQAPAKRLNFGQTDSE 180
DB      121 AKKRVLEPLGLVEGAKTAPGKKRPVEQSPQEBSSSGIOKTDGQAPAKRLNFGQTDSE 180
QY      181 SVDPDPLGHPATTPAAVGTTPAASGGGAPMADNNGAGCGVNASGNWHCDSTWLDGRI 240
DB      181 SVDPDPLGHPATTPAAVGTTPAASGGGAPMADNNGAGCGVNASGNWHCDSTWLDGRI 240
QY      241 TTSRTMALPTYNHLYKQISSASTGASNDNHFGYSTPGYDFPFRFHCFSPRDMQRL 300
DB      241 TTSRTMALPTYNHLYKQISSASTGASNDNHFGYSTPGYDFPFRFHCFSPRDMQRL 300
QY      301 INNNMGFRPKRLNFKLNIQVKEVTNDGVTTIANNLTSTVOVPSDSEYQLPYVLSAQ 360
DB      301 INNNMGFRPKRLNFKLNIQVKEVTNDGVTTIANNLTSTVOVPSDSEYQLPYVLSAQ 360
QY      361 GCLPPPADVFMIPQYGYLTANNNGSQAVERSSFYCLEYFPSONLRTGNNFTFSYTFEDVP 420
DB      361 GCLPPPADVFMIPQYGYLTANNNGSQAVERSSFYCLEYFPSONLRTGNNFTFSYTFEDVP 420
QY      421 FHSYVAHSOSLDRLNMPPLIDQYLYLNRTONQSGAONKDLFRSGSPAGMSYQPKWLP 480
DB      421 FHSYVAHSOSLDRLNMPPLIDQYLYLNRTONQSGAONKDLFRSGSPAGMSYQPKWLP 480
QY      481 GPCYRQORVSKTTDNNNSFTWTGASKYVLRRESIINGGTAMASHKDDKDFPMSGV 540
DB      481 GPCYRQORVSKTTDNNNSFTWTGASKYVLRRESIINGGTAMASHKDDKDFPMSGV 540
QY      541 MIFGKSAGASNTALDNVMTTDEEBEIKATNPVATERFGTAAVNFQSSSTDPATGDVHAMG 600
DB      541 MIFGKSAGASNTALDNVMTTDEEBEIKATNPVATERFGTAAVNFQSSSTDPATGDVHAMG 600
QY      601 ALPGMWQDRDVLQGPIMAKIPHTDGHFSPPLMGFGFKHPPLQILKNTPEVPANPPA 660
DB      601 ALPGMWQDRDVLQGPIMAKIPHTDGHFSPPLMGFGFKHPPLQILKNTPEVPANPPA 660
QY      661 EFSATKFASTITQYSTGQVSEIEMELQKNSKRMPNEVOYTSNYAKSANVDFTVNNGL 720
DB      661 EFSATKFASTITQYSTGQVSEIEMELQKNSKRMPNEVOYTSNYAKSANVDFTVNNGL 720
QY      721 YTEPRPIGTRYLTRPL 736
DB      721 YTEPRPIGTRYLTRPL 736

```

```
RESULT 13
AD227070
ID AD227070 standard; protein; 736 AA.
XX
AC AD227070;
XX
DT 30-JUN-2005 (first entry)
XX
DE Adeno-associated virus protein SEQ ID NO 220.
XX
KM rheumatoid arthritis; multiple sclerosis; sarcoidosis; diabetes;
KM scleroderma; psoriasis; vasculitis; Crohn's disease; hemophilia; HIV;
KM bacterial infection; cancer; ulcerative colitis; antineumatic;
KM antirheumatic; neuroprotective; antiinflammatory; antidiabetic;
KM antipsoiatic; vasotropic; gastrointestinal-gen.; hemostatic; anti-HIV;
KM virucide; antibacterial; cytostatic; antitumor; dermatological.
XX
OS Adeno-associated virus.
XX
PN WO2005033321-A2.
XX
PD 14-APR-2005.
XX
PF 30-SEP-2004; 2004MO-US028817.
XX
PR 30-SEP-2003; 2003US-0508226P.
PR 29-APR-2004; 2004US-0566546P.
XX
PA (UYPE-) UNIV PENNSYLVANIA.
XX
PI Wilson JM, Gao G, Alvira MR, Vandenberghe LH;
XX
DR WPI; 2005-285437/29.
XX
PT New adeno-associated virus (AAV) clade comprising at least three AAV
PT members, useful for preventing and/or treating arthritis, multiple
PT sclerosis, diabetes, scleroderma, psoriasis, hemophilia, HIV, bacterial
PT infection and cancer.
XX
PS Disclosure; SEQ ID NO 220; 569pp; English.
XX
CC The invention relates to an adeno-associated virus (AAV) clade comprising
CC at least three AAV members, where each member of the AAV clade is
CC phylogenetically related as determined using a Neighbor-Joining heuristic
CC by a bootstrap value of at least 75 % per 1000 isolates and a Poisson
CC correction distance measurement of no more than 0.05. The methods and
CC compositions of the present invention are useful for the prevention
CC and/or treatment of rheumatoid arthritis, multiple sclerosis,
CC sarcoidosis, diabetes, scleroderma, psoriasis, vasculitis, Crohn's
CC disease, hemophilia, HIV, bacterial infection, cancer and ulcerative
CC colitis. The present sequence represents the amino acid sequence of an
CC adeno-associated virus protein.
XX
SQ Sequence 736 AA;
Query Match 99.3%; Score 3963; DB 9; Length 736;
Best Local Similarity 99.2%; Pred. No. 6,4e-309;
Matches 730; Conservative 3; Mismatches 3; Indels 0; Gaps 0;
QY 1 MAADGTLPLWLENNLSGIRFEMWDLKPGAPKPYKANOQKODGGLVLPKTYLGPNGLD 60
DB 1 MAADGTLPLWLENNLSGIRFEMWDLKPGAPKPYKANOQKODGGLVLPKTYLGPNGLD 60
QY 61 KGEPVNAADAAAEHKKAYDQOLKAGDNPYLRYNHADAEOERLOEPTSGNUGRAVFO 120
DB 61 KGEPVNAADAAAEHKKAYDQOLKAGDNPYLRYNHADAEOERLOEPTSGNUGRAVFO 120
QY 121 AKRRVLEPIGLVEGAKTAPGKRPVEQSPQEPDSSGGIGTGOQPAKKLINFQGTGSE 180
DB 121 AKRRVLEPIGLVEGAKTAPGKRPVEQSPQEPDSSGGIGTGOQPAKKLINFQGTGSE 180
QY 121 AKRRVLEPIGLVEGAKTAPGKRPVEQSPQEPDSSGGIGTGOQPAKKLINFQGTGSE 180
DB 121 AKRRVLEPIGLVEGAKTAPGKRPVEQSPQEPDSSGGIGTGOQPAKKLINFQGTGSE 180
QY 181 SVDPQPLGEPPTPAVGTTPWASGGGAPMADNNEGADGVGNAGMHCDSITWLGDRVI 240
```

```
DB 181 SVDPQPLGEPPTPAVGTTPWASGGGAPMADNNEGADGVGNAGMHCDSITWLGDRVI 240
QY 241 TTSRTWALPTYNNHLTKQISSASTGASNDNHVFGYSTPMGYPDFNHFHFSBRDQRL 300
DB 241 TTSRTWALPTYNNHLTKQISSASTGASNDNHVFGYSTPMGYPDFNHFHFSBRDQRL 300
QY 301 INNNWGFPRKRLNFKLFNIOVKEVTNDGVTTIANNLTSIVQVFSDEYOLPYVLSAHQ 360
DB 301 INNNWGFPRKRLNFKLFNIOVKEVTNDGVTTIANNLTSIVQVFSDEYOLPYVLSAHQ 360
QY 361 GCLPPPADVFMIPQVGYLLNNGSQAVGRSSFCLEYPFSQMLRTNNFTFSFEFEP 420
DB 361 GCLPPPADVFMIPQVGYLLNNGSQAVGRSSFCLEYPFSQMLRTNNFTFSFEFEP 420
QY 421 FHSYSYASQSLDLMLNPLIDQYLYLNRTONOGSSAQNKLFSRGSFAGMSVOPKXWLP 480
DB 421 FHSYSYASQSLDLMLNPLIDQYLYLNRTONOGSSAQNKLFSRGSFAGMSVOPKXWLP 480
QY 481 GPCYRQORVSKTNDNNNSNFTWTGASKYVNLNGRESIINPGTAMASHKDEDFPMGCV 540
DB 481 GPCYRQORVSKTNDNNNSNFTWTGASKYVNLNGRESIINPGTAMASHKDEDFPMGCV 540
QY 541 MIFGKSAGASNTALDNVMTTDEBEIKATNPVATERGTVAVNFOSSSTDPAQDVHMG 600
DB 541 MIFGKSAGASNTALDNVMTTDEBEIKATNPVATERGTVAVNFOSSSTDPAQDVHMG 600
QY 601 ALPGMWQDRDVLQGPIMAKIPHTDGHFPPSPFAMGFGKPNPQILIKNTVPANPPA 660
DB 601 ALPGMWQDRDVLQGPIMAKIPHTDGHFPPSPFAMGFGKPNPQILIKNTVPANPPA 660
QY 661 EFSATKFASTITQVSTGQVSEIEWELQKENSKRWNDEVOYTSNYAKSANVDFTVNNGL 720
DB 661 EFSATKFASTITQVSTGQVSEIEWELQKENSKRWNDEVOYTSNYAKSANVDFTVNNGL 720
QY 721 YTEPRPIGTRYLTRPL 736
DB 721 YTEPRPIGTRYLTRPL 736
RESULT 14
AD227086
ID AD227086 standard; protein; 737 AA.
XX
AC AD227086;
XX
DT 30-JUN-2005 (first entry)
XX
DE Adeno-associated virus protein SEQ ID NO 236.
XX
KM rheumatoid arthritis; multiple sclerosis; sarcoidosis; diabetes;
KM scleroderma; psoriasis; vasculitis; Crohn's disease; hemophilia; HIV;
KM bacterial infection; cancer; ulcerative colitis; antineumatic;
KM antirheumatic; neuroprotective; antiinflammatory; antidiabetic;
KM antipsoiatic; vasotropic; gastrointestinal-gen.; hemostatic; anti-HIV;
KM virucide; antibacterial; cytostatic; antitumor; dermatological.
XX
OS Adeno-associated virus.
XX
PN WO2005033321-A2.
XX
PD 14-APR-2005.
XX
PF 30-SEP-2004; 2004MO-US028817.
XX
PR 30-SEP-2003; 2003US-0508226P.
PR 29-APR-2004; 2004US-0566546P.
XX
PA (UYPE-) UNIV PENNSYLVANIA.
XX
PI Wilson JM, Gao G, Alvira MR, Vandenberghe LH;
XX
DR WPI; 2005-285437/29.
```

```
XX New adeno-associated virus (AAV) clade comprising at least three AAV
PT members, useful for preventing and/or treating arthritis, multiple
PT sclerosis, diabetes, scleroderma, psoriasis, hemophilia, HIV, bacterial
PT infection and cancer.
XX
PS Claim 19; SEQ ID NO 236; 569pp; English.
XX
CC The invention relates to an adeno-associated virus (AAV) clade comprising
CC at least three AAV members, where each member of the AAV clade is
CC phylogenetically related as determined using a Neighbor-Joining heuristic
CC by a bootstrap value of at least 75 % per 1000 isolates and a Poisson
CC correction distance measurement of no more than 0.05. The methods and
CC compositions of the present invention are useful for the prevention
CC and/or treatment of rheumatoid arthritis, multiple sclerosis,
CC sarcoidosis, diabetes, scleroderma, psoriasis, vasculitis, Crohn's
CC disease, hemophilia, HIV, bacterial infection, cancer and ulcerative
CC colitis. The present sequence represents the amino acid sequence of an
CC adeno-associated virus protein.
XX
SQ Sequence 737 AA;

Query Match          97.7%; Score 3897.5; DB 9; Length 737;
Best Local Similarity 97.7%; Pred. No. 1.2e-303;
Matches 720; Conservative 6; Mismatches 10; Indels 1; Gaps 1;

1 MAADGYLPWLIEDNLSEGIKREWMDLKPGAPKPRANOQKODDGRGLVPGYKYLGPFGNGLD 60
1 MAADGYLPWLIEDNLSEGIKREWMDLKPGAPKPRANOQKODDGRGLVPGYKYLGPFGNGLD 60
61 KGEPVNAADAAALEHDKAYDQOLKAGDNPYLRYNHADAERQERLOEDTSGGNLGRAVFO 120
61 KGEPVNAADAAALEHDKAYDQOLKAGDNPYLRYNHADAERQERLOEDTSGGNLGRAVFO 120
121 AKKRVLEPLGLVEGAKTAPGKRPVEQSPQ-EPDSSSGIGKTGOOPAKKRLNFGOTGDS 179
121 AKKRVLEPLGLVEGAKTAPGKRPVEQSPQ-EPDSSSGIGKTGOOPAKKRLNFGOTGDS 179
121 AKKRVLEPLGLVEGAKTAPGKRPVEQSPQ-EPDSSSGIGKTGOOPAKKRLNFGOTGDS 180
121 AKKRVLEPLGLVEGAKTAPGKRPVEQSPQ-EPDSSSGIGKTGOOPAKKRLNFGOTGDS 180
180 ESYVDPQPLGEPPTPAVAGPTTMAAGGAPMADNNEGADGVNAGNNHCDSTWLGDRV 239
180 ESYVDPQPLGEPPTPAVAGPTTMAAGGAPMADNNEGADGVNAGNNHCDSTWLGDRV 239
181 ESYVDPQPLGEPPTPAVAGPTTMAAGGAPMADNNEGADGVNAGNNHCDSTWLGDRV 240
181 ESYVDPQPLGEPPTPAVAGPTTMAAGGAPMADNNEGADGVNAGNNHCDSTWLGDRV 240
240 ITTSTRTMLPTYNHNLKYKOISASGASNDNHYFGSTPMWGFDPFRFHCHSPDMDOR 239
240 ITTSTRTMLPTYNHNLKYKOISASGASNDNHYFGSTPMWGFDPFRFHCHSPDMDOR 239
241 ITTSTRTMLPTYNHNLKYKOISASGASNDNHYFGSTPMWGFDPFRFHCHSPDMDOR 300
241 ITTSTRTMLPTYNHNLKYKOISASGASNDNHYFGSTPMWGFDPFRFHCHSPDMDOR 300
300 LNNNGGFRPKRLNFKLFIQVKEVTNDGVTIANNLSTVQVFPDSEQLPYVLGSAH 359
300 LNNNGGFRPKRLNFKLFIQVKEVTNDGVTIANNLSTVQVFPDSEQLPYVLGSAH 359
301 LNNNGGFRPKRLNFKLFIQVKEVTNDGVTIANNLSTVQVFPDSEQLPYVLGSAH 360
301 LNNNGGFRPKRLNFKLFIQVKEVTNDGVTIANNLSTVQVFPDSEQLPYVLGSAH 360
360 QGCLPPPADVEMIPOYGYTLTNNGSOAVGRSSPYCLEYFSPQMLRTGNNFTFSYTFEEV 419
360 QGCLPPPADVEMIPOYGYTLTNNGSOAVGRSSPYCLEYFSPQMLRTGNNFTFSYTFEEV 419
361 QGCLPPPADVEMIPOYGYTLTNNGSOAVGRSSPYCLEYFSPQMLRTGNNFTFSYTFEEV 420
361 QGCLPPPADVEMIPOYGYTLTNNGSOAVGRSSPYCLEYFSPQMLRTGNNFTFSYTFEEV 420
420 PFHSSVAHSQSLRLINPLIDQYLYLNRTONSGSQQNDLLPSRSPAGMVSQVOPNNWL 479
420 PFHSSVAHSQSLRLINPLIDQYLYLNRTONSGSQQNDLLPSRSPAGMVSQVOPNNWL 479
421 PLHSSVAHSQSLRLINPLIDQYLYLNRTONSGSQQNDLLPSRSPAGMVSQVOPNNWL 480
421 PLHSSVAHSQSLRLINPLIDQYLYLNRTONSGSQQNDLLPSRSPAGMVSQVOPNNWL 480
480 PGECYRQORVSKTKTDNNNSNFTWTGASKYNLNGRESIIPGTAMASHKDDKFFPMG 539
480 PGECYRQORVSKTKTDNNNSNFTWTGASKYNLNGRESIIPGTAMASHKDDKFFPMG 539
481 PGECYRQORVSKTKTDNNNSNFTWTGASKYNLNGRESIIPGTAMASHKDDKFFPMG 540
481 PGECYRQORVSKTKTDNNNSNFTWTGASKYNLNGRESIIPGTAMASHKDDKFFPMG 540
540 VMI FGKESAGASNTALDNVMTDEEBIKATNPVATERFGTAVANFOSSSTDPAQDVHAM 600
540 VMI FGKESAGASNTALDNVMTDEEBIKATNPVATERFGTAVANFOSSSTDPAQDVHAM 600
541 VMI FGKESAGASNTALDNVMTDEEBIKATNPVATERFGTAVANFOSSSTDPAQDVHAM 600
541 VMI FGKESAGASNTALDNVMTDEEBIKATNPVATERFGTAVANFOSSSTDPAQDVHAM 600
600 GALPGWMDQRDVYLOGPIAKTPTHGHRHPSPLMGFGLLKNPPOILLKNPVPANRP 659
600 GALPGWMDQRDVYLOGPIAKTPTHGHRHPSPLMGFGLLKNPPOILLKNPVPANRP 659
601 GALPGWMDQRDVYLOGPIAKTPTHGHRHPSPLMGFGLLKNPPOILLKNPVPANRP 660
601 GALPGWMDQRDVYLOGPIAKTPTHGHRHPSPLMGFGLLKNPPOILLKNPVPANRP 660
660 AEFSAATFAFITQYSTGVSVLEIEMELQKENSKRANPEVOYTSNYAKSANVDFTVNNG 719
660 AEFSAATFAFITQYSTGVSVLEIEMELQKENSKRANPEVOYTSNYAKSANVDFTVNNG 719
661 AEFSAATFAFITQYSTGVSVLEIEMELQKENSKRANPEVOYTSNYAKSANVDFTVNNG 720
661 AEFSAATFAFITQYSTGVSVLEIEMELQKENSKRANPEVOYTSNYAKSANVDFTVNNG 720
```

```
QY 720 LYTEPRPICTRYLTRPL 736
DB 721 LYTEPRPICTRYLTRPL 737

RESULT 15
ADZ27010
ID ADZ27010 standard; protein; 737 AA.
AC ADZ27010;
XX
XX 30-JUN-2005 (first entry)
XX
XX Adeno-associated virus protein SEQ ID NO 160.
DE
XX
XX rheumatoid arthritis; multiple sclerosis; sarcoidosis; diabetes;
XX scleroderma; psoriasis; vasculitis; Crohn's disease; hemophilia; HIV;
XX bacterial infection; cancer; ulcerative colitis; antineumatic;
XX antirheumatic; neuroprotective; antiinflammatory; antidiabetic;
XX antipsoriatic; vasotrophic; gastrointestinal gen.; hemostatic; anti-HIV;
XX virucide; antibacterial; cytostatic; antiulcer; dermatological.
XX
XX Adeno-associated virus.
OS
XX
XX MO2005033321-A2.
XX
XX 14-APR-2005.
XX
XX 30-SEP-2004; 2004MO-US028817.
XX
XX 30-SEP-2003; 2003US-0508226P.
XX
XX 29-APR-2004; 2004US-0566546P.
XX
XX (UYPE-) UNTIV PENNSYLVANIA.
XX
XX Wilson JM, Gao G, Alvira MR, Vandenbergh LH;
XX
XX WPI; 2005-285437/29.
XX
XX New adeno-associated virus (AAV) clade comprising at least three AAV
XX members, useful for preventing and/or treating arthritis, multiple
XX sclerosis, diabetes, scleroderma, psoriasis, hemophilia, HIV, bacterial
XX infection and cancer.
XX
XX Claim 46; SEQ ID NO 160; 569pp; English.
XX
XX The invention relates to an adeno-associated virus (AAV) clade comprising
XX at least three AAV members, where each member of the AAV clade is
XX phylogenetically related as determined using a Neighbor-Joining heuristic
XX by a bootstrap value of at least 75 % per 1000 isolates and a Poisson
XX correction distance measurement of no more than 0.05. The methods and
XX compositions of the present invention are useful for the prevention
XX and/or treatment of rheumatoid arthritis, multiple sclerosis,
XX sarcoidosis, diabetes, scleroderma, psoriasis, vasculitis, Crohn's
XX disease, hemophilia, HIV, bacterial infection, cancer and ulcerative
XX colitis. The present sequence represents the amino acid sequence of an
XX adeno-associated virus protein.
XX
XX Sequence 737 AA;

Query Match          97.4%; Score 3883.5; DB 9; Length 737;
Best Local Similarity 97.4%; Pred. No. 1.6e-302;
Matches 718; Conservative 6; Mismatches 12; Indels 1; Gaps 1;

1 MAADGYLPWLIEDNLSEGIKREWMDLKPGAPKPRANOQKODDGRGLVPGYKYLGPFGNGLD 60
1 MAADGYLPWLIEDNLSEGIKREWMDLKPGAPKPRANOQKODDGRGLVPGYKYLGPFGNGLD 60
61 KGEPVNAADAAALEHDKAYDQOLKAGDNPYLRYNHADAERQERLOEDTSGGNLGRAVFO 120
61 KGEPVNAADAAALEHDKAYDQOLKAGDNPYLRYNHADAERQERLOEDTSGGNLGRAVFO 120
```


THIS PAGE BLANK (USPTO,

GenCore version 5.1.9
Copyright (c) 1993 - 2006 Biocelebration Ltd.

OM protein - protein search, using sw model

Run on: June 12, 2006, 12:53:30 ; Search time 183 Seconds
(without alignments)
1862.984 Million cell updates/sec

Title: US-10-696-261-13

Perfect score: 3989
Sequence: 1 MADDGYLPDWLEDNLSGIR.....NNGLYTERPRDGTGRYLTPRL 736

Scoring table: BIOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 2097797 seqs, 463214858 residues

Total number of hits satisfying chosen parameters: 2097797

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications_AA Main:*

- 1: /EMC_Celerra_SIDS3/prodata/2/pubppaa/US07_PUBCOMB.pep:*
- 2: /EMC_Celerra_SIDS3/prodata/2/pubppaa/US08_PUBCOMB.pep:*
- 3: /EMC_Celerra_SIDS3/prodata/2/pubppaa/US09_PUBCOMB.pep:*
- 4: /EMC_Celerra_SIDS3/prodata/2/pubppaa/US10_PUBCOMB.pep:*
- 5: /EMC_Celerra_SIDS3/prodata/2/pubppaa/US10_PUBCOMB.pep:*
- 6: /EMC_Celerra_SIDS3/prodata/2/pubppaa/US11_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	3989	100.0	736	US-10-291-583-64	Sequence 64, Appl
2	3989	100.0	736	US-10-423-704A-5	Sequence 5, Appl
3	3989	100.0	736	US-10-696-261-3	Sequence 3, Appl
4	3989	100.0	736	US-10-696-261-13	Sequence 13, Appl
5	3989	100.0	736	US-10-696-282-3	Sequence 3, Appl
6	3989	100.0	736	US-10-696-282-13	Sequence 13, Appl
7	3989	100.0	736	US-10-696-900-3	Sequence 3, Appl
8	3989	100.0	736	US-10-696-900-13	Sequence 13, Appl
9	3989	100.0	736	US-10-496-799-5	Sequence 5, Appl
10	3989	100.0	736	US-10-873-632A-20	Sequence 20, Appl
11	3989	100.0	736	US-11-145-035-20	Sequence 20, Appl
12	3963	99.3	736	US-10-291-583-65	Sequence 65, Appl
13	3963	99.3	736	US-10-873-632A-19	Sequence 19, Appl
14	3963	99.3	736	US-11-145-035-26	Sequence 26, Appl
15	3511	88.0	736	US-11-145-035-22	Sequence 22, Appl
16	3501	87.8	736	US-10-873-632A-18	Sequence 18, Appl
17	3494	87.6	736	US-10-291-583-71	Sequence 71, Appl
18	3494	87.6	736	US-10-423-704A-6	Sequence 6, Appl
19	3494	87.6	736	US-10-496-799-6	Sequence 6, Appl
20	3473	87.1	736	US-10-291-583-97	Sequence 97, Appl
21	3467	86.9	736	US-10-291-583-100	Sequence 100, App
22	3467	86.9	736	US-10-423-704A-7	Sequence 7, Appl
23	3467	86.9	736	US-10-496-799-2	Sequence 2, Appl
24	3460	86.8	736	US-10-291-583-96	Sequence 96, Appl
25	3460	86.7	736	US-10-291-583-98	Sequence 98, Appl
26	3457	86.7	736	US-10-291-583-99	Sequence 99, Appl
27	3453	86.6	738	US-10-291-583-85	Sequence 85, Appl

28	3449	86.5	738	US-11-145-035-41	Sequence 41, Appl
29	3448	86.4	738	US-10-291-583-79	Sequence 79, Appl
30	3448	86.4	738	US-10-291-583-91	Sequence 81, Appl
31	3445	86.4	738	US-10-291-583-93	Sequence 93, Appl
32	3445	86.4	738	US-10-291-583-94	Sequence 94, Appl
33	3442.5	86.3	737	US-10-291-583-2	Sequence 2, Appl
34	3442.5	86.3	737	US-10-423-704A-8	Sequence 8, Appl
35	3442.5	86.3	737	US-11-145-035-38	Sequence 38, Appl
36	3441	86.3	738	US-10-291-583-80	Sequence 80, Appl
37	3440	86.2	738	US-10-291-583-84	Sequence 84, Appl
38	3438	86.2	738	US-10-291-583-92	Sequence 92, Appl
39	3434	86.1	728	US-10-291-583-108	Sequence 108, App
40	3432	86.0	738	US-10-291-583-91	Sequence 91, Appl
41	3429	86.0	738	US-10-291-583-92	Sequence 92, Appl
42	3424.5	85.8	737	US-10-291-583-72	Sequence 82, Appl
43	3422.5	85.8	729	US-10-291-583-110	Sequence 110, App
44	3422	85.8	738	US-10-291-583-83	Sequence 83, Appl
45	3421	85.8	738	US-10-873-632A-21	Sequence 21, Appl

ALIGNMENTS

RESULT 1
US-10-291-583-64
Sequence 64, Application US/10291583
Publication No. US20030138772A1
GENERAL INFORMATION:
APPLICANT: Gao, James M.
APPLICANT: Alviria, Mauricio
TITLE OF INVENTION: A Method of Detecting and/or Identifying Adeno-Associated Virus
FILE REFERENCE: UPN-0273USA
CURRENT APPLICATION NUMBER: US/10/291.583
PRIOR FILING DATE: 2002-11-12
PRIOR APPLICATION NUMBER: US 60/350,607
PRIOR FILING DATE: 2001-11-13
PRIOR APPLICATION NUMBER: US 60/341,117
PRIOR FILING DATE: 2001-12-17
PRIOR APPLICATION NUMBER: US 60/377,066
PRIOR FILING DATE: 2002-05-01
PRIOR APPLICATION NUMBER: US 60/386,675
PRIOR FILING DATE: 2002-06-05
NUMBER OF SEQ ID NOS: 120
SOFTWARE: PatentIn version 3.1
SEQ ID NO 64
LENGTH: 736
TYPE: PRT
ORGANISM: capsid protein of AAV serotype, clone AAV1
US-10-291-583-64
Query Match 100.0%; Score 3989; DB 4; Length 736;
Best Local Similarity 100.0%; Pred. No. 9.5e-309;
Matches 736; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MADDGYLPDWLEDNLSGIREMMDLKPAPKPAKQKODGRLVPGYKYLGPFGNLD 60
DB 1 MADDGYLPDWLEDNLSGIREMMDLKPAPKPAKQKODGRLVPGYKYLGPFGNLD 60
QY 61 KGEPVNAADAALAEHDKAYVQOOLKAGNPLYRINHADAERFQERLQEDTSPFGNIGRAVFQ 120
DB 61 KGEPVNAADAALAEHDKAYVQOOLKAGNPLYRINHADAERFQERLQEDTSPFGNIGRAVFQ 120
QY 121 AKKRVLPELGLVEGATAPGKKRPVQSPQEPSSSGIGTGGQPAKKRLNFGQTDGSE 180
DB 121 AKKRVLPELGLVEGATAPGKKRPVQSPQEPSSSGIGTGGQPAKKRLNFGQTDGSE 180
QY 181 SVDPDPLGPPPTPAVGGTTTMASSGGGAPMADNNEGADVGNAAGNWDSTWLGDRVI 240
DB 181 SVDPDPLGPPPTPAVGGTTTMASSGGGAPMADNNEGADVGNAAGNWDSTWLGDRVI 240
QY 241 TTSTRTVALPTTNNHLTKYKQISSASTGASNDNNHFGISTPMGYPDFNRFCHFSRDMQRL 300

Db	241	TTSTRTMABLPYNNHJLYKQISSASITGASNDNHIFGISTMGJFDFNRFHCHSPROMORL	300
Qy	301	INNMGFRPKRLNFXLFNIQVKEVTTNDGVTTIANNLSTVQVFSDEYQLPVILGSAHQ	360
Db	301	INNMGFRPKRLNFXLFNIQVKEVTTNDGVTTIANNLSTVQVFSDEYQLPVILGSAHQ	360
Qy	361	GCLPPFPADVEMIPOYGYITLNGSGAVGRSSFYCLEYFSPSOMLRGNNFTFSYTEEVP	420
Db	361	GCLPPFPADVEMIPOYGYITLNGSGAVGRSSFYCLEYFSPSOMLRGNNFTFSYTEEVP	420
Qy	421	FHSSVAHSQSLDRMLNPLIDQYLYLNRFQONOSGSAQNODLJFSRQSPGMSVQPKMWLP	480
Db	421	FHSSVAHSQSLDRMLNPLIDQYLYLNRFQONOSGSAQNODLJFSRQSPGMSVQPKMWLP	480
Qy	481	GPCYRQOEVSKTKTNDNNNSNFTWTGASKYNLNGRESIINPGTAMASHKDEDEKFFPMGCV	540
Db	481	GPCYRQOEVSKTKTNDNNNSNFTWTGASKYNLNGRESIINPGTAMASHKDEDEKFFPMGCV	540
Qy	541	MI FGKESAGASNTALDNDWMTDBEEBIKATNPVATEFPGYAVNFOGSSSTDPAATGDYHMG	600
Db	541	MI FGKESAGASNTALDNDWMTDBEEBIKATNPVATEFPGYAVNFOGSSSTDPAATGDYHMG	600
Qy	601	ALPGWVWODRDVYLLGPIWAKI PHTDGHHPSPLMGFGIXNPPQILIKNTEVPANPPA	660
Db	601	ALPGWVWODRDVYLLGPIWAKI PHTDGHHPSPLMGFGIXNPPQILIKNTEVPANPPA	660
Qy	661	EFSATKFPASFITQYSTGVQSVIEIWELOKENSKNMPEVOYTSNVAKSANVDFTVDNNGL	720
Db	661	EFSATKFPASFITQYSTGVQSVIEIWELOKENSKNMPEVOYTSNVAKSANVDFTVDNNGL	720
Qy	721	YTEPRPIGTRYLTRPL 736	
Db	721	YTEPRPIGTRYLTRPL 736	

```

RESULT 2
US-10-423-704A-5
; Sequence 5, Application US/10423704A
; Publication No. US20030228282A1
; GENERAL INFORMATION:
; APPLICANT: Gao, Guangping
; APPLICANT: Wilson, James M.
; APPLICANT: Alvira, Mauricio
; TITLE OF INVENTION: Adeno-Associated Virus (AAV) Serotype 8 Sequences, Vectors
; TITLE OF INVENTION: Containing Same, and Uses Therefor
; FILE REFERENCE: UPN-02733AUSA
; CURRENT APPLICATION NUMBER: US/10/423,704A
; CURRENT FILING DATE: 2003-04-25
; PRIOR APPLICATION NUMBER: US 60/341,151
; PRIOR FILING DATE: 2001-12-17
; PRIOR APPLICATION NUMBER: US 60/377,133
; PRIOR FILING DATE: 2002-05-01
; PRIOR APPLICATION NUMBER: US 60/386,122
; PRIOR FILING DATE: 2002-06-05
; PRIOR APPLICATION NUMBER: PCT/US02/33630
; PRIOR FILING DATE: 2002-11-12
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 5
; LENGTH: 736
; TYPE: PRT
; ORGANISM: adeno-associated virus serotype 1
US-10-423-704A-5

```

Query Match	100.0%	Score 3989	DB 4	Length 736
Best Local Similarity	100.0%	Pred. No. 9.5e-309		
Matches 736	Conservative 0	Mismatches 0	Indels 0	Gaps 0
QY	1	MAADGYLPDMIEDNISEGIRREWMWDIKPAPKPKRANQKODDGRGLVLPGYRYLGPENGLD	60	
Db	1	MAAGYLPDMIEDNISEGIRREWMWDIKPAPKPKRANQKODDGRGLVLPGYRYLGPENGLD	60	

Qy	6	KGEVNAADAALAEHDKAYOOLKAGBNPLR,NHMDAEFOEJLOJDTSGCNLGA	120
Dd	61	KGEVNAADAALAEHDKAYOOLKAGBNPLR,NHMDAEFOEJLOJDTSGCNLGA	120
Qy	121	AKKRVLEPLG,VEBGANTAPGKTRPVEQSPQEDPSSSGIGIKTGOQPAKKRLNF	180
Dd	121	AKKRVLEPLG,VEBGANTAPGKTRPVEQSPQEDPSSSGIGIKTGOQPAKKRLNF	180
Qy	181	SVDPPOJLGPBPATPAVGPBTMASSGGAAMANNAGAVGASGNHCHDSYWLDRYI	240
Dd	181	SVDPPOJLGPBPATPAVGPBTMASSGGAAMANNAGAVGASGNHCHDSYWLDRYI	240
Qy	241	TTSTRYALBPTYNNHLTKYQJSSASTGASDNHFFGYSTWGPEDFNRFCHESPR	300
Dd	241	TTSTRYALBPTYNNHLTKYQJSSASTGASDNHFFGYSTWGPEDFNRFCHESPR	300
Qy	301	INNMGRRPRKLNFLKFLNIQKAVTNTDGYTTIANNLSTVQVPSDSEYQLPV	360
Dd	301	INNMGRRPRKLNFLKFLNIQKAVTNTDGYTTIANNLSTVQVPSDSEYQLPV	360
Qy	361	GCLPFPADYFMI,POXYGLTLNNGSOAVGRSSYCLEYFPSPQMLPTGNNTFS	420
Dd	361	GCLPFPADYFMI,POXYGLTLNNGSOAVGRSSYCLEYFPSPQMLPTGNNTFS	420
Qy	421	FHSSYASOSLDRLMNP,LDQYLYLNRTONOGSAQONKDLFRSGSPAGMSYQ	480
Dd	421	FHSSYASOSLDRLMNP,LDQYLYLNRTONOGSAQONKDLFRSGSPAGMSYQ	480
Qy	481	GPCYRQORVSKTKTDNNNSNFTMGASKYLNLRRESTINGTMAHXODEDF	540
Dd	481	GPCYRQORVSKTKTDNNNSNFTMGASKYLNLRRESTINGTMAHXODEDF	540
Qy	541	MI FCKESAGASNTALDNWMTDEEBEIKANPVATERFGIVAVNFQSSSTD	600
Dd	541	MI FCKESAGASNTALDNWMTDEEBEIKATNPVATERFGIVAVNFQSSSTD	600
Qy	601	ALPBMWODDHYI,OGPIAKIPHTGCHFPSPRLMGFGJXNPPPIILKNTPV	660
Dd	601	ALPBMWODDHYI,OGPIAKIPHTGCHFPSPRLMGFGJXNPPPIILKNTPV	660
Qy	661	EFSATKRASFIITOYSTQGVSEIEMELOKENSKRMNPEVOYTSNVAKSAN	720
Dd	661	EFSATKRASFIITOYSTQGVSEIEMELOKENSKRMNPEVOYTSNVAKSAN	720
Qy	721	YTBPRPIGRTYLRPL 736	
Dd	721	YTBPRPIGRTYLRPL 736	

```

1  RESULT 3
2  US-10-696-261-3
3  / Sequence 3, Application US/10696261
4  / Publication No. US20040057931A1
5  / GENERAL INFORMATION:
6  / APPLICANT: Wilson, James M.
7  / APPLICANT: Xiao, Weidong
8  / TITLE OF INVENTION: Adeno-Associated Virus Serotype I Nucleic Acid Sequences
9  / TITLE OF INVENTION: Vectors and Host Cells Containing Same
10 / FILE REFERENCE: GNVFN.031USA
11 / CURRENT APPLICATION NUMBER: US/10/696,261
12 / CURRENT FILING DATE: 2003-10-29
13 / PRIOR APPLICATION NUMBER: US/09/807,802A
14 / PRIOR FILING DATE: 2002-02-21
15 / PRIOR APPLICATION NUMBER: US 60/107,114
16 / PRIOR FILING DATE: 1998-11-05
17 / PRIOR APPLICATION NUMBER: PCT/US99/25694
18 / PRIOR FILING DATE: 1999-11-02
19 / NUMBER OF SEQ ID NOS: 20
20 / SOFTWARE: PatentIn version 3.1
21 SEQ ID NO 3
22 / LENGTH: 736
23 / TYPE: PRT
24 / ORGANISM: AAV-1

```


US-10-696-261-3

Query Match 100.0%; Score 3989; DB 4; Length 736;
 Best Local Similarity 100.0%; Pred. No. 9.5e-309;
 Matches 736; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

Qy 1 MAAGGYLPDMLEDMLSGIRREWMDLKGAPEKPKANQOKQODDGRGLVLPGYKYLGPFGNGLD 60
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 1 MAAGGYLPDMLEDMLSGIRREWMDLKGAPEKPKANQOKQODDGRGLVLPGYKYLGPFGNGLD 60
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

Qy 61 KGEVYNADAAALAEHDKAYDQQLKAGDNPYLRYNHADAEPQERLQEDTSFGCNLGRAVFQ 120
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 61 KGEVYNADAAALAEHDKAYDQQLKAGDNPYLRYNHADAEPQERLQEDTSFGCNLGRAVFQ 120
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

Qy 121 AKKRVLEPLGLVEGAKTAPGKKRPVEQSPQEPDSSSGIGKTGOOPAKKRLNFGQTGDS 180
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 121 AKKRVLEPLGLVEGAKTAPGKKRPVEQSPQEPDSSSGIGKTGOOPAKKRLNFGQTGDS 180
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

Qy 181 SVDPQPLGEPBPATPAVGPPTTMASGGGAPMADNNEGADGVGNASGNMHCSTWLGDRVI 240
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 181 SVDPQPLGEPBPATPAVGPPTTMASGGGAPMADNNEGADGVGNASGNMHCSTWLGDRVI 240
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

Qy 241 TTSRTMALPTNNHLKQIASSASTGASNDNHFGYSTPMGYPDPNRFCHFSRDMQRL 300
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 241 TTSRTMALPTNNHLKQIASSASTGASNDNHFGYSTPMGYPDPNRFCHFSRDMQRL 300
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

Qy 301 INNNWGRPKRLNFKLFNIQVEYTTNDGVTTIANNLTSTVQVSDSEYQLPYLGSANQ 360
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 301 INNNWGRPKRLNFKLFNIQVEYTTNDGVTTIANNLTSTVQVSDSEYQLPYLGSANQ 360
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

Qy 361 GCLPPFPADVPMIYQGYLTLNNGSOAVGSSFCLEYFSPQMLRTGNFTFSTFEEVP 420
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 361 GCLPPFPADVPMIYQGYLTLNNGSOAVGSSFCLEYFSPQMLRTGNFTFSTFEEVP 420
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

Qy 421 FHSYSYASQSILDRMLNPLIDQYLYLNRTOQSGSAQNKDLLFSRGPAGMSVQPKMLP 480
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 421 FHSYSYASQSILDRMLNPLIDQYLYLNRTOQSGSAQNKDLLFSRGPAGMSVQPKMLP 480
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

Qy 481 GPCYRQORVSKTKTDNNNSNFTWTGASKYNLNGRESIINPGTAMASHKDEDEKFFPMGCV 540
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 481 GPCYRQORVSKTKTDNNNSNFTWTGASKYNLNGRESIINPGTAMASHKDEDEKFFPMGCV 540
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

Qy 541 MIFKESAGASNTALDNVMTDEBEIKATNPVATERFGTAVAVNQSSSTDPAATGDVHAMG 600
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 541 MIFKESAGASNTALDNVMTDEBEIKATNPVATERFGTAVAVNQSSSTDPAATGDVHAMG 600
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

Qy 601 ALPGMWQDRDVLVYQGPIMAKIPHTDGHFHSPLMGFGGLKNPPQILIKNTPVPANPPA 660
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 601 ALPGMWQDRDVLVYQGPIMAKIPHTDGHFHSPLMGFGGLKNPPQILIKNTPVPANPPA 660
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

Qy 661 EFSATKFASTFYQSTQGVSEIEMELQKENSKEWNEVOYTSNYAKSANDFTVDNNGL 720
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 661 EFSATKFASTFYQSTQGVSEIEMELQKENSKEWNEVOYTSNYAKSANDFTVDNNGL 720
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

Qy 721 YTEPRPIGTRYLTRPL 736
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 721 YTEPRPIGTRYLTRPL 736
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

```

RESULT 4

US-10-696-261-13
 ; Sequence 13, Application US/10696261
 ; Publication No. US20040057931A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Wileon, James M.
 ; TITLE OF INVENTION: Adeno-Associated Virus Serotype I Nucleic Acid Sequences,
 ; FILE REFERENCE: GNPVN.031USA
 ; CURRENT APPLICATION NUMBER: US/10/696.261
 ; PRIOR FILING DATE: 2003-10-29
 ; PRIOR APPLICATION NUMBER: US/09/807,802A
 ; PRIOR FILING DATE: 2002-02-21

PRIOR APPLICATION NUMBER: US 60/107,114
 ; PRIOR FILING DATE: 1998-11-05
 ; PRIOR APPLICATION NUMBER: PCT/US99/25694
 ; PRIOR FILING DATE: 1999-11-02
 ; NUMBER OF SEQ ID NOS: 20
 ; SOFTWARE: PatentIn version 3.1
 ; SEQ ID NO 13
 ; LENGTH: 736
 ; TYPE: PRF
 ; ORGANISM: AAV-1
 US-10-696-261-13

Query Match 100.0%; Score 3989; DB 4; Length 736;
 Best Local Similarity 100.0%; Pred. No. 9.5e-309;
 Matches 736; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

Qy 1 MAADGYLPDMLEDMLSGIRREWMDLKGAPEKPKANQOKQODDGRGLVLPGYKYLGPFGNGLD 60
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 1 MAADGYLPDMLEDMLSGIRREWMDLKGAPEKPKANQOKQODDGRGLVLPGYKYLGPFGNGLD 60
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

Qy 61 KGEVYNADAAALAEHDKAYDQQLKAGDNPYLRYNHADAEPQERLQEDTSFGCNLGRAVFQ 120
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 61 KGEVYNADAAALAEHDKAYDQQLKAGDNPYLRYNHADAEPQERLQEDTSFGCNLGRAVFQ 120
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

Qy 121 AKKRVLEPLGLVEGAKTAPGKKRPVEQSPQEPDSSSGIGKTGOOPAKKRLNFGQTGDS 180
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 121 AKKRVLEPLGLVEGAKTAPGKKRPVEQSPQEPDSSSGIGKTGOOPAKKRLNFGQTGDS 180
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

Qy 181 SVDPQPLGEPBPATPAVGPPTTMASGGGAPMADNNEGADGVGNASGNMHCSTWLGDRVI 240
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 181 SVDPQPLGEPBPATPAVGPPTTMASGGGAPMADNNEGADGVGNASGNMHCSTWLGDRVI 240
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

Qy 241 TTSRTMALPTNNHLKQIASSASTGASNDNHFGYSTPMGYPDPNRFCHFSRDMQRL 300
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 241 TTSRTMALPTNNHLKQIASSASTGASNDNHFGYSTPMGYPDPNRFCHFSRDMQRL 300
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

Qy 301 INNNWGRPKRLNFKLFNIQVEYTTNDGVTTIANNLTSTVQVSDSEYQLPYLGSANQ 360
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 301 INNNWGRPKRLNFKLFNIQVEYTTNDGVTTIANNLTSTVQVSDSEYQLPYLGSANQ 360
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

Qy 361 GCLPPFPADVPMIYQGYLTLNNGSOAVGSSFCLEYFSPQMLRTGNFTFSTFEEVP 420
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 361 GCLPPFPADVPMIYQGYLTLNNGSOAVGSSFCLEYFSPQMLRTGNFTFSTFEEVP 420
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

Qy 421 FHSYSYASQSILDRMLNPLIDQYLYLNRTOQSGSAQNKDLLFSRGPAGMSVQPKMLP 480
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 421 FHSYSYASQSILDRMLNPLIDQYLYLNRTOQSGSAQNKDLLFSRGPAGMSVQPKMLP 480
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

Qy 481 GPCYRQORVSKTKTDNNNSNFTWTGASKYNLNGRESIINPGTAMASHKDEDEKFFPMGCV 540
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 481 GPCYRQORVSKTKTDNNNSNFTWTGASKYNLNGRESIINPGTAMASHKDEDEKFFPMGCV 540
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

Qy 541 MIFKESAGASNTALDNVMTDEBEIKATNPVATERFGTAVAVNQSSSTDPAATGDVHAMG 600
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 541 MIFKESAGASNTALDNVMTDEBEIKATNPVATERFGTAVAVNQSSSTDPAATGDVHAMG 600
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

Qy 601 ALPGMWQDRDVLVYQGPIMAKIPHTDGHFHSPLMGFGGLKNPPQILIKNTPVPANPPA 660
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 601 ALPGMWQDRDVLVYQGPIMAKIPHTDGHFHSPLMGFGGLKNPPQILIKNTPVPANPPA 660
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

Qy 661 EFSATKFASTFYQSTQGVSEIEMELQKENSKEWNEVOYTSNYAKSANDFTVDNNGL 720
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 661 EFSATKFASTFYQSTQGVSEIEMELQKENSKEWNEVOYTSNYAKSANDFTVDNNGL 720
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

Qy 721 YTEPRPIGTRYLTRPL 736
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 721 YTEPRPIGTRYLTRPL 736
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

```

RESULT 5

US-10-696-282-3
 ; Sequence 3, Application US/10696282
 ; Publication No. US20040057932A1

```

; GENERAL INFORMATION:
; APPLICANT: Wilson, James M.
; APPLICANT: Xiao, Weidong
; TITLE OF INVENTION: Adeno-Associated Virus Serotype 1 Nucleic Acid Sequences,
; TITLE OF INVENTION: Vectors and Host Cells Containing Same
; FILE REFERENCE: GNPV.031USA
; CURRENT APPLICATION NUMBER: US/10/696,282
; CURRENT FILING DATE: 2003-10-29
; PRIOR APPLICATION NUMBER: US/09/807,802A
; PRIOR FILING DATE: 2002-02-21
; PRIOR APPLICATION NUMBER: US 60/107,114
; PRIOR FILING DATE: 1998-11-05
; PRIOR APPLICATION NUMBER: PCT/US99/25694
; PRIOR FILING DATE: 1999-11-02
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 3
; LENGTH: 736
; TYPE: PRF
; ORGANISM: AAV-1
; US-10-696-282-3

Query Match          100.0%; Score 3989; DB 4; Length 736;
Best Local Similarity 100.0%; Pred. No. 9,5e-309;
Matches 736; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1  MAADGYLPDWLENDLSEGIKREWMWDLKPGAPKPKANQOKDDGRLVPGYKYLGPFGNGLD 60
        |||||||
DB       1  MAADGYLPDWLENDLSEGIKREWMWDLKPGAPKPKANQOKDDGRLVPGYKYLGPFGNGLD 60

QY      61  KGEPVNAADAALAEHDKAYDQOLKAGDNPYLRYNHADAERFOERLQEDTSFGNLSGRAVFQ 120
        |||||||
DB       61  KGEPVNAADAALAEHDKAYDQOLKAGDNPYLRYNHADAERFOERLQEDTSFGNLSGRAVFQ 120

QY      121  AKKRVLEPLGLVEGAKTAPGKKRPVEQSPQEPDSSSGIGKTGOQPAKKRLNFGQTGDS 180
        |||||||
DB       121  AKKRVLEPLGLVEGAKTAPGKKRPVEQSPQEPDSSSGIGKTGOQPAKKRLNFGQTGDS 180

QY      121  AKKRVLEPLGLVEGAKTAPGKKRPVEQSPQEPDSSSGIGKTGOQPAKKRLNFGQTGDS 180
        |||||||
DB       121  AKKRVLEPLGLVEGAKTAPGKKRPVEQSPQEPDSSSGIGKTGOQPAKKRLNFGQTGDS 180

QY      181  SVDPDPLGEPPTPAVGPFTTMASSGGAPMADNNEGADGVGASGNMHCDSYTLGDRV 240
        |||||||
DB       181  SVDPDPLGEPPTPAVGPFTTMASSGGAPMADNNEGADGVGASGNMHCDSYTLGDRV 240

QY      241  TTSRTMALPTYNNHLKYKQISSASTGASNDNHFGYSTPWGYDFNRFCHFSPRDQRL 300
        |||||||
DB       241  TTSRTMALPTYNNHLKYKQISSASTGASNDNHFGYSTPWGYDFNRFCHFSPRDQRL 300

QY      241  TTSRTMALPTYNNHLKYKQISSASTGASNDNHFGYSTPWGYDFNRFCHFSPRDQRL 300
        |||||||
DB       241  TTSRTMALPTYNNHLKYKQISSASTGASNDNHFGYSTPWGYDFNRFCHFSPRDQRL 300

QY      301  INNMGFRPKRLNFKLFNIQVEKVTNDGYTTIANNLSTVOYFSDSEYQLPYVLGSAHQ 360
        |||||||
DB       301  INNMGFRPKRLNFKLFNIQVEKVTNDGYTTIANNLSTVOYFSDSEYQLPYVLGSAHQ 360

QY      301  INNMGFRPKRLNFKLFNIQVEKVTNDGYTTIANNLSTVOYFSDSEYQLPYVLGSAHQ 360
        |||||||
DB       301  INNMGFRPKRLNFKLFNIQVEKVTNDGYTTIANNLSTVOYFSDSEYQLPYVLGSAHQ 360

QY      361  GCLPPPADVFMIPQGYTLTLNNGSOAVGRSSFCLEYPFSOMLRTGNNTFSYTEEVP 420
        |||||||
DB       361  GCLPPPADVFMIPQGYTLTLNNGSOAVGRSSFCLEYPFSOMLRTGNNTFSYTEEVP 420

QY      361  GCLPPPADVFMIPQGYTLTLNNGSOAVGRSSFCLEYPFSOMLRTGNNTFSYTEEVP 420
        |||||||
DB       361  GCLPPPADVFMIPQGYTLTLNNGSOAVGRSSFCLEYPFSOMLRTGNNTFSYTEEVP 420

QY      421  FHSVYAHOSLDBLMPIDQYLYLNRTQNGSAQNKDLFSRGSFAGMSYQPKWLP 480
        |||||||
DB       421  FHSVYAHOSLDBLMPIDQYLYLNRTQNGSAQNKDLFSRGSFAGMSYQPKWLP 480

QY      421  FHSVYAHOSLDBLMPIDQYLYLNRTQNGSAQNKDLFSRGSFAGMSYQPKWLP 480
        |||||||
DB       421  FHSVYAHOSLDBLMPIDQYLYLNRTQNGSAQNKDLFSRGSFAGMSYQPKWLP 480

QY      481  GPCYRQORVSKTIDNNNSFTWTGASKYNLNGRESIINPGTMAHDKDEDFEFPMSV 540
        |||||||
DB       481  GPCYRQORVSKTIDNNNSFTWTGASKYNLNGRESIINPGTMAHDKDEDFEFPMSV 540

QY      541  MIFGKSAGASNTALDNVMTDEBEIKATNPVATERFGTVAVNFQSSSTDPAATGDVHAMG 600
        |||||||
DB       541  MIFGKSAGASNTALDNVMTDEBEIKATNPVATERFGTVAVNFQSSSTDPAATGDVHAMG 600

QY      601  ALPGMWQODDVLQGIWAKIPHTDGHFHSPLMGFGKKNPPQILKNTVPANPAPPA 660
        |||||||
DB       601  ALPGMWQODDVLQGIWAKIPHTDGHFHSPLMGFGKKNPPQILKNTVPANPAPPA 660

QY      601  ALPGMWQODDVLQGIWAKIPHTDGHFHSPLMGFGKKNPPQILKNTVPANPAPPA 660
        |||||||
DB       601  ALPGMWQODDVLQGIWAKIPHTDGHFHSPLMGFGKKNPPQILKNTVPANPAPPA 660

QY      661  EFSATKASFIITQYSTQVSEIEWELQKENSKRMBEVOYTSNYAKSANVDFTVDNGL 720
        |||||||
DB       661  EFSATKASFIITQYSTQVSEIEWELQKENSKRMBEVOYTSNYAKSANVDFTVDNGL 720
```

```

QY      721  YTEPRPIGTRYLTRPL 736
        |||||||
DB       721  YTEPRPIGTRYLTRPL 736

RESULT 6
US-10-696-282-13
; Sequence 13, Application US/10696282
; Publication No. US20040057932A1
; GENERAL INFORMATION:
; APPLICANT: Wilson, James M.
; APPLICANT: Xiao, Weidong
; TITLE OF INVENTION: Adeno-Associated Virus Serotype 1 Nucleic Acid Sequences,
; TITLE OF INVENTION: Vectors and Host Cells Containing Same
; FILE REFERENCE: GNPV.031USA
; CURRENT APPLICATION NUMBER: US/10/696,282
; CURRENT FILING DATE: 2003-10-29
; PRIOR APPLICATION NUMBER: US/09/807,802A
; PRIOR FILING DATE: 2002-02-21
; PRIOR APPLICATION NUMBER: US 60/107,114
; PRIOR FILING DATE: 1998-11-05
; PRIOR APPLICATION NUMBER: PCT/US99/25694
; PRIOR FILING DATE: 1999-11-02
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 13
; LENGTH: 736
; TYPE: PRF
; ORGANISM: AAV-1
; US-10-696-282-13

Query Match          100.0%; Score 3989; DB 4; Length 736;
Best Local Similarity 100.0%; Pred. No. 9,5e-309;
Matches 736; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1  MAADGYLPDWLENDLSEGIKREWMWDLKPGAPKPKANQOKDDGRLVPGYKYLGPFGNGLD 60
        |||||||
DB       1  MAADGYLPDWLENDLSEGIKREWMWDLKPGAPKPKANQOKDDGRLVPGYKYLGPFGNGLD 60

QY      61  KGEPVNAADAALAEHDKAYDQOLKAGDNPYLRYNHADAERFOERLQEDTSFGNLSGRAVFQ 120
        |||||||
DB       61  KGEPVNAADAALAEHDKAYDQOLKAGDNPYLRYNHADAERFOERLQEDTSFGNLSGRAVFQ 120

QY      121  AKKRVLEPLGLVEGAKTAPGKKRPVEQSPQEPDSSSGIGKTGOQPAKKRLNFGQTGDS 180
        |||||||
DB       121  AKKRVLEPLGLVEGAKTAPGKKRPVEQSPQEPDSSSGIGKTGOQPAKKRLNFGQTGDS 180

QY      121  AKKRVLEPLGLVEGAKTAPGKKRPVEQSPQEPDSSSGIGKTGOQPAKKRLNFGQTGDS 180
        |||||||
DB       121  AKKRVLEPLGLVEGAKTAPGKKRPVEQSPQEPDSSSGIGKTGOQPAKKRLNFGQTGDS 180

QY      181  SVDPDPLGEPPTPAVGPFTTMASSGGAPMADNNEGADGVGASGNMHCDSYTLGDRV 240
        |||||||
DB       181  SVDPDPLGEPPTPAVGPFTTMASSGGAPMADNNEGADGVGASGNMHCDSYTLGDRV 240

QY      181  SVDPDPLGEPPTPAVGPFTTMASSGGAPMADNNEGADGVGASGNMHCDSYTLGDRV 240
        |||||||
DB       181  SVDPDPLGEPPTPAVGPFTTMASSGGAPMADNNEGADGVGASGNMHCDSYTLGDRV 240

QY      241  TTSRTMALPTYNNHLKYKQISSASTGASNDNHFGYSTPWGYDFNRFCHFSPRDQRL 300
        |||||||
DB       241  TTSRTMALPTYNNHLKYKQISSASTGASNDNHFGYSTPWGYDFNRFCHFSPRDQRL 300

QY      241  TTSRTMALPTYNNHLKYKQISSASTGASNDNHFGYSTPWGYDFNRFCHFSPRDQRL 300
        |||||||
DB       241  TTSRTMALPTYNNHLKYKQISSASTGASNDNHFGYSTPWGYDFNRFCHFSPRDQRL 300

QY      301  INNMGFRPKRLNFKLFNIQVEKVTNDGYTTIANNLSTVOYFSDSEYQLPYVLGSAHQ 360
        |||||||
DB       301  INNMGFRPKRLNFKLFNIQVEKVTNDGYTTIANNLSTVOYFSDSEYQLPYVLGSAHQ 360

QY      301  INNMGFRPKRLNFKLFNIQVEKVTNDGYTTIANNLSTVOYFSDSEYQLPYVLGSAHQ 360
        |||||||
DB       301  INNMGFRPKRLNFKLFNIQVEKVTNDGYTTIANNLSTVOYFSDSEYQLPYVLGSAHQ 360

QY      361  GCLPPPADVFMIPQGYTLTLNNGSOAVGRSSFCLEYPFSOMLRTGNNTFSYTEEVP 420
        |||||||
DB       361  GCLPPPADVFMIPQGYTLTLNNGSOAVGRSSFCLEYPFSOMLRTGNNTFSYTEEVP 420

QY      361  GCLPPPADVFMIPQGYTLTLNNGSOAVGRSSFCLEYPFSOMLRTGNNTFSYTEEVP 420
        |||||||
DB       361  GCLPPPADVFMIPQGYTLTLNNGSOAVGRSSFCLEYPFSOMLRTGNNTFSYTEEVP 420

QY      421  FHSVYAHOSLDBLMPIDQYLYLNRTQNGSAQNKDLFSRGSFAGMSYQPKWLP 480
        |||||||
DB       421  FHSVYAHOSLDBLMPIDQYLYLNRTQNGSAQNKDLFSRGSFAGMSYQPKWLP 480

QY      421  FHSVYAHOSLDBLMPIDQYLYLNRTQNGSAQNKDLFSRGSFAGMSYQPKWLP 480
        |||||||
DB       421  FHSVYAHOSLDBLMPIDQYLYLNRTQNGSAQNKDLFSRGSFAGMSYQPKWLP 480

QY      481  GPCYRQORVSKTIDNNNSFTWTGASKYNLNGRESIINPGTMAHDKDEDFEFPMSV 540
        |||||||
DB       481  GPCYRQORVSKTIDNNNSFTWTGASKYNLNGRESIINPGTMAHDKDEDFEFPMSV 540

QY      481  GPCYRQORVSKTIDNNNSFTWTGASKYNLNGRESIINPGTMAHDKDEDFEFPMSV 540
        |||||||
DB       481  GPCYRQORVSKTIDNNNSFTWTGASKYNLNGRESIINPGTMAHDKDEDFEFPMSV 540

QY      541  MIFGKSAGASNTALDNVMTDEBEIKATNPVATERFGTVAVNFQSSSTDPAATGDVHAMG 600
        |||||||
DB       541  MIFGKSAGASNTALDNVMTDEBEIKATNPVATERFGTVAVNFQSSSTDPAATGDVHAMG 600
```

```
Db 541 MIFGESAGSANTLADNMWITDEBEIKATNPVATERFCTVAVNFQSSSTDATGVAHMG 600
Qy 601 ALPGWVWQDRDVLVYLGQPIWAKI PHTDGHFHPSPLMGFGGLKNPPQILIKXTVPVAPNPA 660
Db 601 ALPGWVWQDRDVLVYLGQPIWAKI PHTDGHFHPSPLMGFGGLKNPPQILIKXTVPVAPNPA 660
Qy 661 EFSATKPFASFTQYSTQGVSVIEIWELOKENSKRMPVEQYTSNVAKSANVDFTVDNNGL 720
Db 661 EFSATKPFASFTQYSTQGVSVIEIWELOKENSKRMPVEQYTSNVAKSANVDFTVDNNGL 720
Qy 721 YTEPRPIGTRYLTRPL 736
Db 721 YTEPRPIGTRYLTRPL 736

RESULT 7
US-10-696-900-3
; Sequence 3, Application US/10696900
; Publication No. US20040057933A1
; GENERAL INFORMATION:
; APPLICANT: Wilson, James M.
; APPLICANT: Xiao, Weidong
; TITLE OF INVENTION: Adeno-Associated Virus Serotype 1 Nucleic Acid Sequences,
; FILE OF INVENTION: Vectors and Host Cells Containing Same
; FILE REFERENCE: GNPVN.031USA
; CURRENT APPLICATION NUMBER: US/10/696,900
; PRIOR FILING DATE: 2003-10-30
; PRIOR APPLICATION NUMBER: US/09/807,802A
; PRIOR FILING DATE: 2002-02-21
; PRIOR APPLICATION NUMBER: US 60/107,114
; PRIOR FILING DATE: 1998-11-05
; PRIOR APPLICATION NUMBER: PCT/US99/25694
; PRIOR FILING DATE: 1999-11-02
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 3
; LENGTH: 736
; TYPE: PRT
; ORGANISM: AAV-1
US-10-696-900-3

Query Match 100.0%; Score 3989; DB 4; Length 736;
Best Local Similarity 100.0%; Pred. No. 9.5e-309;
Matches 736; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MAADGYLPDMLENDLSGIREMMDLKGPAPKPKANQOKODDGRGLVLPGYKYLGPFGNLD 60
Db 1 MAADGYLPDMLENDLSGIREMMDLKGPAPKPKANQOKODDGRGLVLPGYKYLGPFGNLD 60
Qy 61 KGEFVNADAALAEHDKAYDQOLKAGDNPYLRYNHADAEPERLQEDTSPFGNIGRAVFQ 120
Db 61 KGEFVNADAALAEHDKAYDQOLKAGDNPYLRYNHADAEPERLQEDTSPFGNIGRAVFQ 120
Qy 121 AKKRVLEPLGLVEGAKTAPGKKRPVBSPOEPDSSSGIGTGQOPAKKRLNFGQTDSE 180
Db 121 AKKRVLEPLGLVEGAKTAPGKKRPVBSPOEPDSSSGIGTGQOPAKKRLNFGQTDSE 180
Qy 121 AKKRVLEPLGLVEGAKTAPGKKRPVBSPOEPDSSSGIGTGQOPAKKRLNFGQTDSE 180
Db 121 AKKRVLEPLGLVEGAKTAPGKKRPVBSPOEPDSSSGIGTGQOPAKKRLNFGQTDSE 180
Qy 181 SVDPPOPLGEBPPATPAVGPFTTMASSGGAAPADNNEGADGVGNAAGMHCDSITWLGDRVI 240
Db 181 SVDPPOPLGEBPPATPAVGPFTTMASSGGAAPADNNEGADGVGNAAGMHCDSITWLGDRVI 240
Qy 241 TTSRTWALPTYNHLYKQISSASTGASNDNHFGYSTPMGYPDFNRFCHFSRDMQRL 300
Db 241 TTSRTWALPTYNHLYKQISSASTGASNDNHFGYSTPMGYPDFNRFCHFSRDMQRL 300
Qy 301 INNMGRRPRKRLNKLRFIYQKEYTTNDGVTIANNTLSTYQVSDSYQLPYVLGSAHQ 360
Db 301 INNMGRRPRKRLNKLRFIYQKEYTTNDGVTIANNTLSTYQVSDSYQLPYVLGSAHQ 360
Qy 361 GCLPPPADVPMIYOYGYLTNNNGSOAVGRSSFYCLETFPSQMLRTGNNPFSTTFEEVP 420
Db 361 GCLPPPADVPMIYOYGYLTNNNGSOAVGRSSFYCLETFPSQMLRTGNNPFSTTFEEVP 420
```

```
Qy 421 FHSYSVHQSQSIDRLMNPILIDQYLYYLNRTONQSGASQNKDILFSGRSPAGMSVOPKMWLP 480
Db 421 FHSYSVHQSQSIDRLMNPILIDQYLYYLNRTONQSGASQNKDILFSGRSPAGMSVOPKMWLP 480
Qy 481 GPCYRQORVSKTKDNNNSNFTWVGASKYVNLNGRESIINPCTAMASHDODDKPFPMSGV 540
Db 481 GPCYRQORVSKTKDNNNSNFTWVGASKYVNLNGRESIINPCTAMASHDODDKPFPMSGV 540
Qy 541 MIFGESAGSANTLADNMWITDEBEIKATNPVATERFCTVAVNFQSSSTDATGVAHMG 600
Db 541 MIFGESAGSANTLADNMWITDEBEIKATNPVATERFCTVAVNFQSSSTDATGVAHMG 600
Qy 601 ALPGWVWQDRDVLVYLGQPIWAKI PHTDGHFHPSPLMGFGGLKNPPQILIKXTVPVAPNPA 660
Db 601 ALPGWVWQDRDVLVYLGQPIWAKI PHTDGHFHPSPLMGFGGLKNPPQILIKXTVPVAPNPA 660
Qy 661 EFSATKPFASFTQYSTQGVSVIEIWELOKENSKRMPVEQYTSNVAKSANVDFTVDNNGL 720
Db 661 EFSATKPFASFTQYSTQGVSVIEIWELOKENSKRMPVEQYTSNVAKSANVDFTVDNNGL 720
Qy 721 YTEPRPIGTRYLTRPL 736
Db 721 YTEPRPIGTRYLTRPL 736

RESULT 8
US-10-696-900-13
; Sequence 13, Application US/10696900
; Publication No. US20040057933A1
; GENERAL INFORMATION:
; APPLICANT: Wilson, James M.
; APPLICANT: Xiao, Weidong
; TITLE OF INVENTION: Adeno-Associated Virus Serotype 1 Nucleic Acid Sequences,
; FILE OF INVENTION: Vectors and Host Cells Containing Same
; FILE REFERENCE: GNPVN.031USA
; CURRENT APPLICATION NUMBER: US/10/696,900
; PRIOR FILING DATE: 2003-10-30
; PRIOR APPLICATION NUMBER: US/09/807,802A
; PRIOR FILING DATE: 2002-02-21
; PRIOR APPLICATION NUMBER: US 60/107,114
; PRIOR FILING DATE: 1998-11-05
; PRIOR APPLICATION NUMBER: PCT/US99/25694
; PRIOR FILING DATE: 1999-11-02
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 13
; LENGTH: 736
; TYPE: PRT
; ORGANISM: AAV-1
US-10-696-900-13

Query Match 100.0%; Score 3989; DB 4; Length 736;
Best Local Similarity 100.0%; Pred. No. 9.5e-309;
Matches 736; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MAADGYLPDMLENDLSGIREMMDLKGPAPKPKANQOKODDGRGLVLPGYKYLGPFGNLD 60
Db 1 MAADGYLPDMLENDLSGIREMMDLKGPAPKPKANQOKODDGRGLVLPGYKYLGPFGNLD 60
Qy 61 KGEFVNADAALAEHDKAYDQOLKAGDNPYLRYNHADAEPERLQEDTSPFGNIGRAVFQ 120
Db 61 KGEFVNADAALAEHDKAYDQOLKAGDNPYLRYNHADAEPERLQEDTSPFGNIGRAVFQ 120
Qy 121 AKKRVLEPLGLVEGAKTAPGKKRPVBSPOEPDSSSGIGTGQOPAKKRLNFGQTDSE 180
Db 121 AKKRVLEPLGLVEGAKTAPGKKRPVBSPOEPDSSSGIGTGQOPAKKRLNFGQTDSE 180
Qy 181 SVDPPOPLGEBPPATPAVGPFTTMASSGGAAPADNNEGADGVGNAAGMHCDSITWLGDRVI 240
Db 181 SVDPPOPLGEBPPATPAVGPFTTMASSGGAAPADNNEGADGVGNAAGMHCDSITWLGDRVI 240
Qy 241 TTSRTWALPTYNHLYKQISSASTGASNDNHFGYSTPMGYPDFNRFCHFSRDMQRL 300
```



```
Query Match      100.0%; Score 3989; DB 5; Length 736;
Best Local Similarity 100.0%; Pred. No. 9.5e-309;
Matches 736; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 MAADGYLPTDMLDNLSEGIKREWMWDLKPGAPKPKANQOKQODDGRGLVLPGYKYLGPFGNLD 60
Db 1 MAADGYLPTDMLDNLSEGIKREWMWDLKPGAPKPKANQOKQODDGRGLVLPGYKYLGPFGNLD 60
Oy 61 KGEVYNADAAALAEHDKAYDQOLKAGDNPLYRYNHADAPEORLOEDTSFGSGLGRAVFO 120
Db 61 KGEVYNADAAALAEHDKAYDQOLKAGDNPLYRYNHADAPEORLOEDTSFGSGLGRAVFO 120
Oy 61 KGEVYNADAAALAEHDKAYDQOLKAGDNPLYRYNHADAPEORLOEDTSFGSGLGRAVFO 120
Db 61 KGEVYNADAAALAEHDKAYDQOLKAGDNPLYRYNHADAPEORLOEDTSFGSGLGRAVFO 120
Oy 121 AKKAVLEPLGLVEGAKTAPGKKRPVEOSPOEPDSSSGIGTKGOOPAKKRLNFGQTGSE 180
Db 121 AKKAVLEPLGLVEGAKTAPGKKRPVEOSPOEPDSSSGIGTKGOOPAKKRLNFGQTGSE 180
Oy 121 AKKAVLEPLGLVEGAKTAPGKKRPVEOSPOEPDSSSGIGTKGOOPAKKRLNFGQTGSE 180
Db 121 AKKAVLEPLGLVEGAKTAPGKKRPVEOSPOEPDSSSGIGTKGOOPAKKRLNFGQTGSE 180
Oy 181 SVDPPOPUGEPBPATPAAGPTTMASSGGAPMADNNEGADGVGNASGNMHCDSTWLGDRVI 240
Db 181 SVDPPOPUGEPBPATPAAGPTTMASSGGAPMADNNEGADGVGNASGNMHCDSTWLGDRVI 240
Oy 181 SVDPPOPUGEPBPATPAAGPTTMASSGGAPMADNNEGADGVGNASGNMHCDSTWLGDRVI 240
Db 181 SVDPPOPUGEPBPATPAAGPTTMASSGGAPMADNNEGADGVGNASGNMHCDSTWLGDRVI 240
Oy 241 TTSTRTALPTYNHLYKQISSASTGASNDNHYFGYSTPMGYFDPNRFCHFSPPDMOQL 300
Db 241 TTSTRTALPTYNHLYKQISSASTGASNDNHYFGYSTPMGYFDPNRFCHFSPPDMOQL 300
Oy 301 INNNWGRPKRLNFKLFNIQVKEVTNDGVTTIANNLSTYQVPSDSSEYOLPYVLGSAHQ 360
Db 301 INNNWGRPKRLNFKLFNIQVKEVTNDGVTTIANNLSTYQVPSDSSEYOLPYVLGSAHQ 360
Oy 301 INNNWGRPKRLNFKLFNIQVKEVTNDGVTTIANNLSTYQVPSDSSEYOLPYVLGSAHQ 360
Db 301 INNNWGRPKRLNFKLFNIQVKEVTNDGVTTIANNLSTYQVPSDSSEYOLPYVLGSAHQ 360
Oy 361 GCLPPPADVPMIOYGYLTNNNGSOAVGRSSFCLEYFSPSOMLRTGNPFPSYFEEVP 420
Db 361 GCLPPPADVPMIOYGYLTNNNGSOAVGRSSFCLEYFSPSOMLRTGNPFPSYFEEVP 420
Oy 361 GCLPPPADVPMIOYGYLTNNNGSOAVGRSSFCLEYFSPSOMLRTGNPFPSYFEEVP 420
Db 361 GCLPPPADVPMIOYGYLTNNNGSOAVGRSSFCLEYFSPSOMLRTGNPFPSYFEEVP 420
Oy 421 FHSSVYASQSILDRMLNPLIDQYLYLNRTQOQSSAQNKDLFSGSGPAGMSVOPKMWLP 480
Db 421 FHSSVYASQSILDRMLNPLIDQYLYLNRTQOQSSAQNKDLFSGSGPAGMSVOPKMWLP 480
Oy 421 FHSSVYASQSILDRMLNPLIDQYLYLNRTQOQSSAQNKDLFSGSGPAGMSVOPKMWLP 480
Db 421 FHSSVYASQSILDRMLNPLIDQYLYLNRTQOQSSAQNKDLFSGSGPAGMSVOPKMWLP 480
Oy 481 GPCYRQORVSKTKTDNNNSNFTWTGASKYKYNLNGRESIINPGTAMASHKODEDKFFPMGCV 540
Db 481 GPCYRQORVSKTKTDNNNSNFTWTGASKYKYNLNGRESIINPGTAMASHKODEDKFFPMGCV 540
Oy 481 GPCYRQORVSKTKTDNNNSNFTWTGASKYKYNLNGRESIINPGTAMASHKODEDKFFPMGCV 540
Db 481 GPCYRQORVSKTKTDNNNSNFTWTGASKYKYNLNGRESIINPGTAMASHKODEDKFFPMGCV 540
Oy 541 MIFKESAGASNTALDNVMTDEBEIKATNPVATERFCTVAVNFQSSSTDPAATGVDVHMG 600
Db 541 MIFKESAGASNTALDNVMTDEBEIKATNPVATERFCTVAVNFQSSSTDPAATGVDVHMG 600
Oy 541 MIFKESAGASNTALDNVMTDEBEIKATNPVATERFCTVAVNFQSSSTDPAATGVDVHMG 600
Db 541 MIFKESAGASNTALDNVMTDEBEIKATNPVATERFCTVAVNFQSSSTDPAATGVDVHMG 600
Oy 601 ALPGMWQODRDVYLGQPIWAKIPHTDGHFHSPLMGFGLKNPPQILIKTTPVAPNPPA 660
Db 601 ALPGMWQODRDVYLGQPIWAKIPHTDGHFHSPLMGFGLKNPPQILIKTTPVAPNPPA 660
Oy 601 ALPGMWQODRDVYLGQPIWAKIPHTDGHFHSPLMGFGLKNPPQILIKTTPVAPNPPA 660
Db 601 ALPGMWQODRDVYLGQPIWAKIPHTDGHFHSPLMGFGLKNPPQILIKTTPVAPNPPA 660
Oy 661 EFSATKFPASFTYQSTQGVSEIEMELQKENSCKMNPDEVQYTSNVAKSANVDFVNNGL 720
Db 661 EFSATKFPASFTYQSTQGVSEIEMELQKENSCKMNPDEVQYTSNVAKSANVDFVNNGL 720
Oy 721 YTEBRPIGTRYLTRPL 736
Db 721 YTEBRPIGTRYLTRPL 736

RESULT 11
US-11-145-035-20
; Sequence 20, Application US/11145035
; Publication No. US20050287122A1
; GENERAL INFORMATION:
; APPLICANT: Bartlett et al.
; TITLE OF INVENTION: AAV VECTORS AND METHODS
; FILE REFERENCE: 28335/41335
; CURRENT APPLICATION NUMBER: US/11/145.035
; PRIOR FILING DATE: 2005-06-03
; PRIOR APPLICATION NUMBER: US 10/038,972
; PRIOR FILING DATE: 2002-01-04
; PRIOR APPLICATION NUMBER: US 60/260,124
; NUMBER OF SEQ ID NOS: 45
; SOFTWARE: PatentIn version 3.3
```

```
; SEQ ID NO 20
; LENGTH: 736
; TYPE: PRT
; ORGANISM: Adeno-associated virus 1
US-11-145-035-20

Query Match      100.0%; Score 3989; DB 6; Length 736;
Best Local Similarity 100.0%; Pred. No. 9.5e-309;
Matches 736; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 MAADGYLPTDMLDNLSEGIKREWMWDLKPGAPKPKANQOKQODDGRGLVLPGYKYLGPFGNLD 60
Db 1 MAADGYLPTDMLDNLSEGIKREWMWDLKPGAPKPKANQOKQODDGRGLVLPGYKYLGPFGNLD 60
Oy 61 KGEVYNADAAALAEHDKAYDQOLKAGDNPLYRYNHADAPEORLOEDTSFGSGLGRAVFO 120
Db 61 KGEVYNADAAALAEHDKAYDQOLKAGDNPLYRYNHADAPEORLOEDTSFGSGLGRAVFO 120
Oy 61 KGEVYNADAAALAEHDKAYDQOLKAGDNPLYRYNHADAPEORLOEDTSFGSGLGRAVFO 120
Db 61 KGEVYNADAAALAEHDKAYDQOLKAGDNPLYRYNHADAPEORLOEDTSFGSGLGRAVFO 120
Oy 121 AKKAVLEPLGLVEGAKTAPGKKRPVEOSPOEPDSSSGIGTKGOOPAKKRLNFGQTGSE 180
Db 121 AKKAVLEPLGLVEGAKTAPGKKRPVEOSPOEPDSSSGIGTKGOOPAKKRLNFGQTGSE 180
Oy 121 AKKAVLEPLGLVEGAKTAPGKKRPVEOSPOEPDSSSGIGTKGOOPAKKRLNFGQTGSE 180
Db 121 AKKAVLEPLGLVEGAKTAPGKKRPVEOSPOEPDSSSGIGTKGOOPAKKRLNFGQTGSE 180
Oy 181 SVDPPOPUGEPBPATPAAGPTTMASSGGAPMADNNEGADGVGNASGNMHCDSTWLGDRVI 240
Db 181 SVDPPOPUGEPBPATPAAGPTTMASSGGAPMADNNEGADGVGNASGNMHCDSTWLGDRVI 240
Oy 181 SVDPPOPUGEPBPATPAAGPTTMASSGGAPMADNNEGADGVGNASGNMHCDSTWLGDRVI 240
Db 181 SVDPPOPUGEPBPATPAAGPTTMASSGGAPMADNNEGADGVGNASGNMHCDSTWLGDRVI 240
Oy 241 TTSTRTALPTYNHLYKQISSASTGASNDNHYFGYSTPMGYFDPNRFCHFSPPDMOQL 300
Db 241 TTSTRTALPTYNHLYKQISSASTGASNDNHYFGYSTPMGYFDPNRFCHFSPPDMOQL 300
Oy 301 INNNWGRPKRLNFKLFNIQVKEVTNDGVTTIANNLSTYQVPSDSSEYOLPYVLGSAHQ 360
Db 301 INNNWGRPKRLNFKLFNIQVKEVTNDGVTTIANNLSTYQVPSDSSEYOLPYVLGSAHQ 360
Oy 301 INNNWGRPKRLNFKLFNIQVKEVTNDGVTTIANNLSTYQVPSDSSEYOLPYVLGSAHQ 360
Db 301 INNNWGRPKRLNFKLFNIQVKEVTNDGVTTIANNLSTYQVPSDSSEYOLPYVLGSAHQ 360
Oy 361 GCLPPPADVPMIOYGYLTNNNGSOAVGRSSFCLEYFSPSOMLRTGNPFPSYFEEVP 420
Db 361 GCLPPPADVPMIOYGYLTNNNGSOAVGRSSFCLEYFSPSOMLRTGNPFPSYFEEVP 420
Oy 361 GCLPPPADVPMIOYGYLTNNNGSOAVGRSSFCLEYFSPSOMLRTGNPFPSYFEEVP 420
Db 361 GCLPPPADVPMIOYGYLTNNNGSOAVGRSSFCLEYFSPSOMLRTGNPFPSYFEEVP 420
Oy 421 FHSSVYASQSILDRMLNPLIDQYLYLNRTQOQSSAQNKDLFSGSGPAGMSVOPKMWLP 480
Db 421 FHSSVYASQSILDRMLNPLIDQYLYLNRTQOQSSAQNKDLFSGSGPAGMSVOPKMWLP 480
Oy 421 FHSSVYASQSILDRMLNPLIDQYLYLNRTQOQSSAQNKDLFSGSGPAGMSVOPKMWLP 480
Db 421 FHSSVYASQSILDRMLNPLIDQYLYLNRTQOQSSAQNKDLFSGSGPAGMSVOPKMWLP 480
Oy 481 GPCYRQORVSKTKTDNNNSNFTWTGASKYKYNLNGRESIINPGTAMASHKODEDKFFPMGCV 540
Db 481 GPCYRQORVSKTKTDNNNSNFTWTGASKYKYNLNGRESIINPGTAMASHKODEDKFFPMGCV 540
Oy 481 GPCYRQORVSKTKTDNNNSNFTWTGASKYKYNLNGRESIINPGTAMASHKODEDKFFPMGCV 540
Db 481 GPCYRQORVSKTKTDNNNSNFTWTGASKYKYNLNGRESIINPGTAMASHKODEDKFFPMGCV 540
Oy 541 MIFKESAGASNTALDNVMTDEBEIKATNPVATERFCTVAVNFQSSSTDPAATGVDVHMG 600
Db 541 MIFKESAGASNTALDNVMTDEBEIKATNPVATERFCTVAVNFQSSSTDPAATGVDVHMG 600
Oy 541 MIFKESAGASNTALDNVMTDEBEIKATNPVATERFCTVAVNFQSSSTDPAATGVDVHMG 600
Db 541 MIFKESAGASNTALDNVMTDEBEIKATNPVATERFCTVAVNFQSSSTDPAATGVDVHMG 600
Oy 601 ALPGMWQODRDVYLGQPIWAKIPHTDGHFHSPLMGFGLKNPPQILIKTTPVAPNPPA 660
Db 601 ALPGMWQODRDVYLGQPIWAKIPHTDGHFHSPLMGFGLKNPPQILIKTTPVAPNPPA 660
Oy 601 ALPGMWQODRDVYLGQPIWAKIPHTDGHFHSPLMGFGLKNPPQILIKTTPVAPNPPA 660
Db 601 ALPGMWQODRDVYLGQPIWAKIPHTDGHFHSPLMGFGLKNPPQILIKTTPVAPNPPA 660
Oy 661 EFSATKFPASFTYQSTQGVSEIEMELQKENSCKMNPDEVQYTSNVAKSANVDFVNNGL 720
Db 661 EFSATKFPASFTYQSTQGVSEIEMELQKENSCKMNPDEVQYTSNVAKSANVDFVNNGL 720
Oy 721 YTEBRPIGTRYLTRPL 736
Db 721 YTEBRPIGTRYLTRPL 736

RESULT 12
US-10-291-583-65
; Sequence 65, Application US/10291583
; Publication No. US20030138772A1
; GENERAL INFORMATION:
; APPLICANT: Gao, Guangping
; APPLICANT: Wilson, James M.
; APPLICANT: Alvir, Maurice O.
; TITLE OF INVENTION: A Method of Detecting and/or Identifying Adeno-Associated Virus
; Sequences and Isolating No. US20030138772A1 Sequences Identif
```

FILE REFERENCE: UPN-02735USA
CURRENT APPLICATION NUMBER: US/10/291,583
CURRENT FILING DATE: 2002-11-12
PRIOR APPLICATION NUMBER: US 60/350,607
PRIOR FILING DATE: 2001-11-13
PRIOR APPLICATION NUMBER: US 60/341,117
PRIOR FILING DATE: 2001-12-17
PRIOR APPLICATION NUMBER: US 60/377,066
PRIOR FILING DATE: 2002-05-01
PRIOR APPLICATION NUMBER: US 60/386,675
PRIOR FILING DATE: 2002-06-05
NUMBER OF SEQ ID NOS: 120
SOFTWARE: PatentIn version 3.1
SEQ ID NO 65
LENGTH: 736
TYPE: PRT
ORGANISM: capsid protein of AAV serotype, clone AAV6VP1
US-10-291-583-65

Query Match 99.3%; Score 3963; DB 4; Length 736;
Best Local Similarity 99.2%; Pred. No. 1.le-306;
Matches 730; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

1 MAADGTLPMLEBNDLSEGIKREWMDLKPGAPKPKANOOKODDGRGLVPGYKYLGPFGD 60
1 MAADGTLPMLEBNDLSEGIKREWMDLKPGAPKPKANOOKODDGRGLVPGYKYLGPFGD 60
61 KGEPPVNAADAAALEHDKAVDQOLKAGDNPYLRYNHADAERFOERLQEDTSFGNLGRAVFO 120
61 KGEPPVNAADAAALEHDKAVDQOLKAGDNPYLRYNHADAERFOERLQEDTSFGNLGRAVFO 120
121 AKKRVLEPLGLVEGAKTAPGKKRPVEQSPQEPDSSSGIGKTQOQPAKKRLNFGQTGDS 180
121 AKKRVLEPLGLVEGAKTAPGKKRPVEQSPQEPDSSSGIGKTQOQPAKKRLNFGQTGDS 180
121 AKKRVLEPLGLVEGAKTAPGKKRPVEQSPQEPDSSSGIGKTQOQPAKKRLNFGQTGDS 180
181 SVDDPOLGEPPTPAVGPPTWASGGGAPMADNNEGADGVNAGSNWCHDSITWLGDRVI 240
181 SVDDPOLGEPPTPAVGPPTWASGGGAPMADNNEGADGVNAGSNWCHDSITWLGDRVI 240
241 TTSRTWALPTYNHLYKQISSASTGASNDNHFGYSTPMGYDFPNRPHCHFSPRDQRL 300
241 TTSRTWALPTYNHLYKQISSASTGASNDNHFGYSTPMGYDFPNRPHCHFSPRDQRL 300
301 INNMGFRPKRLNFKLFNIQVKEVTINDGVTIANNLSTVQVFSDEYOLPYVLGSAHQ 360
301 INNMGFRPKRLNFKLFNIQVKEVTINDGVTIANNLSTVQVFSDEYOLPYVLGSAHQ 360
301 INNMGFRPKRLNFKLFNIQVKEVTINDGVTIANNLSTVQVFSDEYOLPYVLGSAHQ 360
361 GCLPPPADVFMIPQGYLTLLNNGSOAVGRSSFCLEYPFSOMLRGNNTFTSTYEDVP 420
361 GCLPPPADVFMIPQGYLTLLNNGSOAVGRSSFCLEYPFSOMLRGNNTFTSTYEDVP 420
421 FHSYAHOSGLDLMLNPLIDQYLYLNRTQNGSAGNOKDLFSRGS PAGMSVOPKMWLP 480
421 FHSYAHOSGLDLMLNPLIDQYLYLNRTQNGSAGNOKDLFSRGS PAGMSVOPKMWLP 480
421 FHSYAHOSGLDLMLNPLIDQYLYLNRTQNGSAGNOKDLFSRGS PAGMSVOPKMWLP 480
481 GPCYRQORVSKITDNNNSFTWTGASKYNLNGRESIINPGTAMASHKDDKDFPFMSGV 540
481 GPCYRQORVSKITDNNNSFTWTGASKYNLNGRESIINPGTAMASHKDDKDFPFMSGV 540
541 MIGKESAGASNTALDNVMTTDEEIKATNPVATERFGTVAVNFQSSSTDPATGDVHMG 600
541 MIGKESAGASNTALDNVMTTDEEIKATNPVATERFGTVAVNFQSSSTDPATGDVHMG 600
601 ALPGMWQDRDYLLOGPIMAKIPHTDGHFSPMLMGFGKHPRPQILLIKNTIPVAPNPA 660
601 ALPGMWQDRDYLLOGPIMAKIPHTDGHFSPMLMGFGKHPRPQILLIKNTIPVAPNPA 660
661 EFSATKASFTIOYSTQOVSEIEMELQKENSKRANBEVOYTSYASAVDFVNDGL 720
661 EFSATKASFTIOYSTQOVSEIEMELQKENSKRANBEVOYTSYASAVDFVNDGL 720
721 YTEBRPIGTRYLTRPL 736
721 YTEBRPIGTRYLTRPL 736

Db 721 YTEBRPIGTRYLTRPL 736

RESULT 13
US-10-873-632A-19
Sequence 19, Application US/10873632A
Publication No. US2006005133A1
GENERAL INFORMATION:
APPLICANT: ARBETMAN, Alejandra E
APPLICANT: COLOSI, Peter
APPLICANT: LOCHRIE, Michael A.
APPLICANT: SUROSKY, Richard T.
TITLE OF INVENTION: AAV VIRIONS WITH DECREASED IMMUNOREACTIVITY AND USES THEREFOR
FILE REFERENCE: 0800-0032
CURRENT APPLICATION NUMBER: US/10/873,632A
CURRENT FILING DATE: 2004-06-21
PRIOR APPLICATION NUMBER: 60/480,395
PRIOR FILING DATE: 2003-06-19
PRIOR APPLICATION NUMBER: 60/567,310
PRIOR FILING DATE: 2004-04-30
PRIOR APPLICATION NUMBER: 60/576,501
PRIOR FILING DATE: 2004-06-03
NUMBER OF SEQ ID NOS: 28
SOFTWARE: PatentIn version 3.3
SEQ ID NO 19
LENGTH: 736
TYPE: PRT
ORGANISM: adeno-associated virus 6
US-10-873-632A-19

Query Match 99.3%; Score 3963; DB 5; Length 736;
Best Local Similarity 99.2%; Pred. No. 1.le-306;
Matches 730; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

1 MAADGTLPMLEBNDLSEGIKREWMDLKPGAPKPKANOOKODDGRGLVPGYKYLGPFGD 60
1 MAADGTLPMLEBNDLSEGIKREWMDLKPGAPKPKANOOKODDGRGLVPGYKYLGPFGD 60
61 KGEPPVNAADAAALEHDKAVDQOLKAGDNPYLRYNHADAERFOERLQEDTSFGNLGRAVFO 120
61 KGEPPVNAADAAALEHDKAVDQOLKAGDNPYLRYNHADAERFOERLQEDTSFGNLGRAVFO 120
121 AKKRVLEPLGLVEGAKTAPGKKRPVEQSPQEPDSSSGIGKTQOQPAKKRLNFGQTGDS 180
121 AKKRVLEPLGLVEGAKTAPGKKRPVEQSPQEPDSSSGIGKTQOQPAKKRLNFGQTGDS 180
121 AKKRVLEPLGLVEGAKTAPGKKRPVEQSPQEPDSSSGIGKTQOQPAKKRLNFGQTGDS 180
181 SVDDPOLGEPPTPAVGPPTWASGGGAPMADNNEGADGVNAGSNWCHDSITWLGDRVI 240
181 SVDDPOLGEPPTPAVGPPTWASGGGAPMADNNEGADGVNAGSNWCHDSITWLGDRVI 240
241 TTSRTWALPTYNHLYKQISSASTGASNDNHFGYSTPMGYDFPNRPHCHFSPRDQRL 300
241 TTSRTWALPTYNHLYKQISSASTGASNDNHFGYSTPMGYDFPNRPHCHFSPRDQRL 300
301 INNMGFRPKRLNFKLFNIQVKEVTINDGVTIANNLSTVQVFSDEYOLPYVLGSAHQ 360
301 INNMGFRPKRLNFKLFNIQVKEVTINDGVTIANNLSTVQVFSDEYOLPYVLGSAHQ 360
301 INNMGFRPKRLNFKLFNIQVKEVTINDGVTIANNLSTVQVFSDEYOLPYVLGSAHQ 360
361 GCLPPPADVFMIPQGYLTLLNNGSOAVGRSSFCLEYPFSOMLRGNNTFTSTYEDVP 420
361 GCLPPPADVFMIPQGYLTLLNNGSOAVGRSSFCLEYPFSOMLRGNNTFTSTYEDVP 420
421 FHSYAHOSGLDLMLNPLIDQYLYLNRTQNGSAGNOKDLFSRGS PAGMSVOPKMWLP 480
421 FHSYAHOSGLDLMLNPLIDQYLYLNRTQNGSAGNOKDLFSRGS PAGMSVOPKMWLP 480
421 FHSYAHOSGLDLMLNPLIDQYLYLNRTQNGSAGNOKDLFSRGS PAGMSVOPKMWLP 480
481 GPCYRQORVSKITDNNNSFTWTGASKYNLNGRESIINPGTAMASHKDDKDFPFMSGV 540
481 GPCYRQORVSKITDNNNSFTWTGASKYNLNGRESIINPGTAMASHKDDKDFPFMSGV 540
541 MIGKESAGASNTALDNVMTTDEEIKATNPVATERFGTVAVNFQSSSTDPATGDVHMG 600
541 MIGKESAGASNTALDNVMTTDEEIKATNPVATERFGTVAVNFQSSSTDPATGDVHMG 600


```

Qy 421 FHSSYAHSGSLDRMLNPLIDQYLYLNRIQ-NOSGSAONKOLLFSRGSFPAKMSVQPRNML 479
Db 420 FHSSYAHSGSLDRMLNPLIDQYLYLNRIQGTTSGETTOSRLLFSQAGPQSMISLOARNML 479
Qy 480 PGPCYRQOORVSKTKTNNNSNFMTGTGASKYNLNGRESIINPGTAMASHKDEDEKFFPMG 539
Db 480 PGPCYRQOORVSKTKTNNNSNFMTGTGASKYNLNGRESIINPGTAMASHKDEDEKFFPMG 539
Qy 540 VMIFGKESAGASNTALDNWMITDEBEIKATNPVATERFGTVAVNFQSSSTDPAIGDVHAM 599
Db 540 NLI FGKGTASNAEIDNWMITDEBEIRITNPVATEQGTVANLQSSNTAPTTRTVNDQ 599
Qy 600 GALPGWVWQDRDYVYLOGPTWAKIPTHGHHFSPMLMGFGLKNPPQIILIKNTVPANPP 659
Db 600 GALPGWVWQDRDYVYLOGPTWAKIPTHGHHFSPMLMGFGLKNPPQIILIKNTVPANPP 659
Qy 660 AEFSAATKFAFTQYSTGQVSEIEMELOKENSRRNPEVOYTSNYAKSANVDFTVDNNG 719
Db 660 TTFSPAKFASFTQYSTGQVSEIEMELOKENSRRNPEIQTYSNINXSVNVDFTVDNG 719
Qy 720 LYTEPRPIGTRVYLRPL 736
Db 720 VYSEPRPIGTRVYLRNL 736

```

Search completed: June 12, 2006, 12:57:21
 Job time : 185 secs

GenCore version 5.1.9
Copyright (c) 1993 - 2006 Bioacceleration Ltd.

OM protein - protein search, using sw model

Run on: June 12, 2006, 12:47:29 ; Search time 45 Seconds
(without alignments)
1573.679 Million cell updates/sec

Title: US-10-696-261-13

Perfect score: 3989
Sequence: 1 MADDGYPDWLEDNLSGIR.....NNGLYTEPRPIGRYTRPL 736

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR 80: *
1: pir1: *
2: pir2: *
3: pir3: *
4: pir4: *

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2177.5	54.6	504	1 VCPV3A	coat protein - ade
2	2168.5	54.4	732	2 SS2210	coat protein Vp1 -
3	798.5	20.0	673	1 VCPV85	coat protein Vp1 -
4	600.5	15.1	781	1 VCPV19	coat protein Vp1 -
5	482	12.1	729	1 A60006	coat protein Vp1 -
6	482	12.1	729	1 VCPVNA	coat protein Vp1 -
7	444.5	11.1	727	1 VCPV1F	coat protein Vp1 -
8	439.5	11.0	722	1 VCPVME	coat protein Vp1 -
9	431.5	10.8	727	1 VCPVFP	coat protein Vp1 -
10	426	10.7	723	1 VCPVPP	coat protein Vp1 -
11	424.5	10.6	718	1 VCPVIM	coat protein Vp1 -
12	421.5	10.6	748	1 VCPVCP	coat protein Vp1 -
13	407.5	10.2	737	1 VCPVCD	coat protein Vp1 -
14	404	10.1	722	1 VCPVVC	coat protein Vp1 -
15	385.5	9.7	722	1 VCPVCN	coat protein Vp1 -
16	377	9.5	716	1 VCPV2M	coat protein Vp1 -
17	234	5.9	587	1 B44276	coat protein Vp1 -
18	199.5	5.0	810	2 A44054	orf1 protein - Jun
19	198	5.0	584	2 S49594	capsid protein VP2
20	191	4.8	702	1 VCPVAP	coat protein Vp1 -
21	143.5	3.6	648	2 S50856	whn protein - rat
22	139	3.5	931	2 T49710	related to glucan
23	131.5	3.3	1446	1 A45344	immediate-early pr
24	130.5	3.3	1145	2 T18235	transcription acti
25	130	3.3	1338	2 T30565	MAP kinase kinase
26	129	3.2	1072	2 A86827	hypothetical prote
27	124.5	3.1	2493	2 A55481	adenylate cyclase
28	124	3.1	1296	2 C62521	hemolysin-type cal
29	124	3.1	3078	2 T28432	variant-specific s

30	123.5	3.1	690	2 AB0124	probable TonB-depe
31	123.5	3.1	2529	2 B64635	toxlin-like outer m
32	123	3.1	1142	2 T37455	enaminein precursor
33	122.5	3.1	880	1 SYBSVS	valine-tRNA ligase
34	121	3.0	635	2 F96660	nuclear P2X1.10 (
35	121	3.0	1113	2 S28925	nuclear pore compl
36	121	3.0	1549	2 T48103	type VII collagen
37	120	3.0	1379	2 T37752	hypothetical serin
38	119.5	3.0	1180	2 E86719	hypothetical prote
39	119.5	3.0	1394	2 S60762	IGA-specific serin
40	119.5	3.0	2004	2 AC0314	probable membrane
41	119	3.0	667	2 A41311	transcription fact
42	118.5	3.0	2399	2 H1879	toxlin-like outer m
43	118	3.0	739	2 T52289	probable transkeo
44	118	3.0	956	2 T08144	myrosinase-binding
45	117.5	2.9	1076	1 A35622	nuclear pore prote

ALIGNMENTS

RESULT 1					
VCPV3A					
coat protein - adeno-associated virus type 2					
C:Species: adeno-associated virus type 2					
C:Date: 05-Apr-1983 #sequence_revision 05-Apr-1983 #text_change 09-Jul-2004					
C:Accession: A03698					
R:Srivastava, A.; Lueby, E.W.; Berns, K.I.					
J. Virol. 45, 555-564, 1983					
A:Title: Nucleotide sequence and organization of the adeno-associated virus 2 genome.					
A:Reference number: A03694; MUID:83164299; PMID:6300419					
A:Accession: A03698					
A:Status: translation not shown					
A:Molecule type: DNA					
A:Residues: 1-504 <SRI>					
A:Cross-references: UNIPROT:P03135; UNIPARC:UPI0000127C8E; EMBL:J01901; NID:g209616; PR					
C:Superfamily: adeno-associated virus coat protein					
C:Keywords: coat protein					
Query Match	54.6%	Score 2177.5	DB 1	Length 504	
Best Local Similarity	82.5%	Pred. No. 3.9e-134			
Matches 392	Conservative 33	Mismatches 47	Indels 3	Gaps 2	
QY	203	MASGGAPMADNNEGADGAGNAGNMGHCDSTWLGDRVITTSRTWALPTNNHLYKQISS	262		
DB	1	MATSGAPMADNNEGADGAGNAGNMGHCDSTWMDRVITTSRTWALPTNNHLYKQISS	60		
QY	263	ASTGASNDNHFFGYSTWGYDFENRPHCHSPRDWQRLNNMGFRPRRLNFKLFNIQVK	322		
DB	61	QS-GASNDNHFFGYSTWGYDFENRPHCHSPRDWQRLNNMGFRPRRLNFKLFNIQVK	119		
QY	323	EVTNDGQTTIANNLTSTVQVFSDEYQLPYVLSAAGQCLPPPADVFMIPQGYLTIN	382		
DB	120	EVTNDGQTTIANNLTSTVQVFTDSYQLPYVLSAAGQCLPPPADVFMIPQGYLTIN	179		
QY	383	NGSQAVRRSSFFCYCLEYPPSQMLRTGNFTFSYFEEVPHSSVYASOSGLRLNPLIDQY	442		
DB	180	NGSQAVRRSSFFCYCLEYPPSQMLRTGNFTFSYFEEVPHSSVYASOSGLRLNPLIDQY	239		
QY	443	LYYNRTQONGSAGNNDLLFSRSGPAGMSQPKNWLPGPCYRQORVSKTTDNNNSNFT	502		
DB	240	LYYNRTQNGSAGNNDLLFSRSGPAGMSQPKNWLPGPCYRQORVSKTTDNNNSNFT	299		
QY	503	WTGASKYNLNGRESIINPGTAMASHKDDKDFPMSGVMIFGKESAGASNTALDNVMTD	562		
DB	300	WTGATKYNLNGRDLNVP--AMASHKDDKDFPMSGVMIFGKESAGASNTALDNVMTD	357		
QY	563	EEETKATNPVATTEFGYVAVNFGSSSTDPAIGDVHAGALPGMWQODRDVYLOGPYAKI	622		
DB	358	EEETGTTNPVATTEFGYVAVNFGSSSTDPAIGDVHAGALPGMWQODRDVYLOGPYAKI	417		
QY	623	PHTGHPSPSLMGFGGLKPPPOILIKNTPVNPANPAEFATKFAPIITQYSTG	677		

QY 636 GGFGLKNPPPOILK--NTVPANPPEAFSATKASPTTQYSTQVSEIEMELQKENS 693
D 577 GSIMSHSPPTGIFIKLARIIPVGNQD-----SFLNIVYTCQVSCVVEWEKRGTK 627
QY 694 RMNEVOVTSNYAKSAND-FTVDNNGLY 721
D 628 NMRPEYMS--ATNMSVDATYINNAGVY 653

RESULT 4
VCPV19
coat protein VP1 - parvovirus B19 (strain Au)
C/Species: parvovirus B19
A/Note: host Homo sapiens (man)
C/Date: 28-Dec-1987 #sequence_revision 28-Dec-1987 #text_change 09-Jul-2004
C/Accession: A24299
R/Shade, R.O.; Blundell, M.C.; Cotmore, S.F.; Tattersall, P.; Astell, C.R.
J. Virol. 58, 921-936, 1986
A/Title: Nucleotide sequence and genome organization of human parvovirus B19 isolated fr
A/Reference number: A24299; MUID:86200451; PMID:3701931
A/Accession: A24299
A/Molecule type: DNA
A/Residues: 1-781 <SHA>
A/Cross-references: UNIPROT:P07299; UNIPARC:UPI000012707C; EMBL:M13178; NID:G333375; PID
C/Superfamily: parvovirus coat protein
C/Keywords: coat protein

Query Match 15.1%; Score 600.5; DB 1; Length 781;
Best Local Similarity 26.6%; Pred. No. 3.5e-31;
Matches 182; Conservative 95; Mismatches 300; Indels 107; Gaps 20;

QY 47 LPGYKYLGPFGNGDKGEFVNAADAALAHDKAVDOQLKAGDNPYLRYNHADAEPOLQ 106
D 125 LPGNYVYGPENELQAGPPQSAVDSARHIDFRYGLKLGINPYTHMTVADBELLNKIKN 184
QY 107 DTSGGVLGRAVFOAKKRVLEPLGLVEGAKTAPGKKRPVQSPQEPDSSSGIGTKGQP 166
D 185 ETGFQAVVVDXVF-----TLKGAAPVVAHFQ-----GSLPEVP 217
QY 167 AKKRLNGQGTGDSVDPDPLGEPPTPAVGPPTTMASSGGGAPMADNEGADGVNAGS 226
D 218 AYNA-----SKYPSMTSVNSAEASTGA-----GGG-----GSNSKXS 250
QY 227 NMHCDSTWLGDRVITSTRTALPTVNNHLYKQISSASTGASND-----NHFGY 276
D 251 MMSEGAFFSANSVCTFSROFLIYDPEHNHYKVPSPASSCHMASGKEAKYCTISPLMGY 310
QY 277 STPWGYFDENRFCHFSFRDQRLINNNGFRPKRLNFKLFIQVKEVT--TNDGVTTIA 334
D 311 STPRYIDFVALNMFPSPLFEQHLIENYGSIAFPDALVTTISEIAVKDVTGTGGV-QVT 369
QY 335 NNLIISTVOVSDSEYOLPYVLGSAHQGLPFPADVFMIPQYGLTLNN--GSAVG---- 389
D 370 DSTTGRICMLVDHEYKPYVLGGQODTLAPFLPTWVFPQYALITVGDVNTQGISGDSK 429
QY 390 -----RSSFYCLEFSPQMLRTGNMFTFSYFEEVPRHSSYAHOSQIDRLNPLIDQYLY 444
D 430 KLASSEAFYVLBEHSSPQLLGTGTASMSYKFPVPRENLEGGQHFEWYNPL---YGS 486
QY 445 YLNRTONQSSAQNKDLFFSRGSPAGMSVQPKMLPQPCYRQQRVSKTKTDNNNSNTWT 504
D 487 RLGVFDTLGGDPKFRSL-----THEDHAIQPNFTMPGVLVNSVSTKGGDSNTAGKALT 541
QY 505 GASKYNNLNGRESIINPG-TMAASHKDEDEKFPMSGVWIPKESAGASNTLNDV----- 558
D 542 GLSTGTSQNTRISLRPGVSPQRYTHMTDKYVTGIMNISHGQTTYG--NADDEXYQGV 598
QY 559 --MITDEEIKATNPVATERGTVAVNFQSSSTDPATGVDYAMGALPQMWQDRVYLOG 616
D 599 GRFNEKEQLKQGLMMHTY-----FPNKGTOQYTDQIE-RPLAMGVSNNRRLALHYES 651
QY 617 PIMAKIHTDGHFHS-PLMGGFLKNPPPOILKNTPVANPPEAFSATKASPTTQYST 675

D 652 QLMSEKIPVLDSEFTQFPAALGQWGLHDPPOIPLK--ILFQSGPIGIGKSMGITTLVQYA 709
QY 676 TGOVSEIEIEMEL-OKENSKRNPE 698
D 710 VGIMTVMTFTLGGPRKATGRWNPQ 733

RESULT 5
A60006
coat protein VP1 - porcine parvovirus (strain 90HS)
N/Contains: coat protein VP2
C/Species: porcine parvovirus
C/Date: 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 09-Jul-2004
C/Accession: A60006
R/Sakurai, M.; Niehmoori, T.; Ushimi, C.; Nakajima, H.
Virus Res. 13, 79-86, 1989
A/Title: Nucleotide sequence of capsid protein gene of porcine parvovirus.
A/Reference number: A60006; MUID:85319168; PMID:2750278
A/Accession: A60006
A/Molecule type: DNA
A/Residues: 1-729 <SAK>
A/Cross-references: UNIPROT:P33484; UNIPARC:UPI0000127C6E
C/Superfamily: parvovirus coat protein
C/Keywords: coat protein; glycoprotein
F/151-729/Product: coat protein VP2 #status predicted <VP2>
F/172,198,282,330,433,471,573,604,651/Binding site: carbohydrate (Asn) (covalent) #stat

Query Match 12.1%; Score 482; DB 1; Length 729;
Best Local Similarity 25.7%; Pred. No. 1.7e-23;
Matches 197; Conservative 112; Mismatches 314; Indels 144; Gaps 35;

QY 43 RGLVLPGYKYLGPFGNGDKGEFVNAADAALAHDKAVDOQLKAGDNPYLRYNHADAEPOL 102
D 9 RGLTLPGYKYLGPENSLDQGEPTNPDAAKEHDAVDKYKSGNPYFYSADDEKRTX 68
QY 103 RLQEDTSPGNGLGRAVFOAKKRVLEPLGLVEGAKTAPGKKRPVQSPQEPDSSSGIGK- 161
D 69 ETEHAKDYGKGIGHYFPRRA-KRAFRP--KLESTOSPPTSQQPEVRSPRKHGSPPKR 125
QY 162 -----TGOQPAKKRLNCGQGTGDSVDPDPLGEPPTPAVGPPTTMASSGGGAPMA 212
D 126 PAPRHITNLAKKRAKKTSTNNSMSBENQHPNINAGTEL-SATGESGGGGG-- 181
QY 213 DNNEGADGVNAGSNMWH--CDSTWLGCD--RVITSTRTALPTVNNHLYKQIS--SAST 265
D 182 -GGRGAGGVGVSSTGSPNNQTEFOYLGEGIVRITMAASRLHLNPREHTYKRIHVANS 240
QY 266 GAS---NDNHFGYSTPWPGYFDFNRFCHFSFRDQRLINNNGFRPKRLNFKLFIQV 321
D 241 GVAQQWQODPAHTQMTVPMSLIDANANGVWFNPADOLISNMTEINLVSPFOEIPNVVL 300
QY 322 KEVT---TNDGVTTIANLIISTVOVFSDEXQLPYVLGSAHQGLPFPADVFMIPQYGL 378
D 301 KLTIESATSPPTKYNNDLASTLAWALDYNNTLTTPPAAPSESLTGLGYMLPTPEYQRY 360
QY 379 L-----TLNNGSOAVGRS-----SFYCLE-YFPQMLRTGNMF-TFSYTE 417
D 361 YLSTCRMLNPTTYGSGQOQITDSIQTGLHSDIMFYTENAVPHILRTGDFSGIYHFD 420
QY 418 EVFPHSSYAHOSQIDR-----LAMPPLI--DOYLYLNRTONQSSAQNKDLFFSRGS 467
D 421 TKPL--KLTHSWQNRSLGLRPKLLTPEETEGDHPGTLPAANRKGQHTMNSYTEAT 478
QY 468 PAGMSVQPKMLPQPCYRQQRVSKTKTDNNNSNTWTGASKYNNLNGRESIINP--GTMA 525
D 479 ---AIRP-----KQVYNTPRANFEYNGGPF-----LTPVLPADT 512
QY 526 SHKDEDEKFPMSGVWIPKESAGASNTALDNWMTDEEIKATNP-----VATERGT 579
D 513 QYNDDE-----PNGAIRTWGQHQGLTTS-----SQELERYTENPQSKGRARXQDPNQ 562
QY 580 VA-VNFQSSS-----TDPAITG--DYAMGALP-----GNVWQDRDYVLOGRPAKIPHT 625

```

Db      563 QAPLNLNTNNNGTLPSDPLGGKPNMFMMTLNTYGPLTALNNTAPVPFNGQIWDEKELDT 622
Qy      626 D--GHFHSPLMGFGFLK-NPPQILIKNTVPANPPAEFSA-TKFASFITYOSTGVSV 681
Db      623 DLKPRH---VTAPFVCCKNPPGQLFVKIAP---NLTDPMADSPOQPRITITSNFMWK 676
Qy      682 ELEMELOKENSKRMNPEVOYTSNYSKANSANVDFTVDNNGLYTEPRPIG 728
Db      677 TLFTFAKMRSSNMNMNFIQHTT-----TAENIGNYI-PTNIG 712

RESULT 6
coat protein VP1 - porcine parvovirus (strain NADL-2)
N:Contains: coat protein VP2
C:Species: porcine parvovirus
C>Date: 31-Dec-1990 #sequence_revision 31-Dec-1990 #text_change 09-Jul-2004
C:Accession: B33743; D48472
R:Vasudevacharya, U.; Basak, S.; Srinivas, R.V.; Compans, R.W.
Virology 173, 368-377, 1989
A>Title: Nucleotide sequence analysis of the capsid genes and the right-hand terminal pa
A:Reference number: A33743; MUID:90085785; PMID:2596019
A:Accession: B33743
A:Molecule type: DNA
A:Residues: 1-729 <VAS>
A:Cross-references: UNIPROT:P18546; UNIPARC:UPI0000127C70; GB:M32787; NID:G332983; PIDN:
R:Bergeron, J.; Meneses, J.; Tjissen, P.
Virology 197, 86-98, 1993
A>Title: Genomic organization and mapping of transcription and translation products of p
A:Reference number: A48472; MUID:94025614; PMID:8212558
A:Accession: D48472
A:Molecule type: DNA
A:Residues: 11-729 <BER>
A:Cross-references: UNIPARC:UPI0000174965
A:Experimental source: NADL-2, ATCC VR-742
A>Note: sequence extracted from NCBI backbone (NCBIN:138789, NCBI:P138794)
C:Genetics:
A:Intons: 10/1
C:Superfamily: parvovirus coat protein
C:Keywords: coat protein; glycoprotein
F:151-729/Product: coat protein VP2 #status predicted <VP2>
F:12,172,198,282,330,433,441,573,604,651/Binding site: carbohydrate (asn) (covalent) #st

Query Match      12.1%; Score 482; DB 1; Length 729;
Best Local Similarity 25.5%; Pred. No. 1,7e-23;
Matches 197; Conservative 111; Mismatches 310; Indels 154; Gaps 35;

Qy      43 RGLVLPGYKYLGPFGNGLDKGEPVNADAALEHDKAVDQOLKAGDNFYLYRNHADAFOE 102
Db      9 RGLTLPGYKYLGPNSLSDGEPTNPDAAKEHDEADKYIKSGKNFYFYSAADKFIK 68
Qy      103 RLQEDTSPFGNLGRAVFAQAKRVLEPLGLVEBGAKTAPGKKRPEVDSPOEDSSSGIGK 161
Db      69 ETEHADYGGKIGHYFFRAKARAPKL---SETDSPITSQOPEYRARRPRKHGSKPRGKR 125
Qy      162 -----TGQOPAKKRLNFGQTGSESVDPDQPLGPPATPAVGPPTTMAAGGAPMA 212
Db      126 PADRHIFINLAKKAKGTSNTNSMSSENVQHNPIAGTEL-SATNESGGGGGGS--- 181
Qy      213 DNNEGADVGNAAG--NMHCDSTWLGD---RVITTSRTMALPPTNNHLYKQIS--SAST 265
Db      182 -GGRGAGGVGVSCTGFNNQTEFOYLGGLVRLTAHAARLIHLNMPHEETKRIHVLSSES 240
Qy      266 GAS-----NDMHYFGYSTPMGYPFDFNRFCHFSPPDMQRLINNMGFRPKRLNFKLFINIQ 321
Db      241 GVAQGVQDDAHQWTFPMSLIDANAGWFPNPDMDLISNMTEILVAFEGEIRFVVL 300
Qy      322 KEVT---TNDGVTTIANLTSTVQVFSDESEYQLPYVLGSAHQGLPPFPADEVMIPOGY 378
Db      301 KITTESATSPPTKLYNNDLTASLWALDTNNLTLPYTPAAPSESTLGFYPMLPKPTGYRY 360
Qy      379 L-----TLNNGSQAVGRS-----STYCLE-YTPSQMLATGNNF-TFSITYE 417
```

```

Db      361 YLSCIRNLNPEYTGSGSQQTIDSIGTGLSDIMFYILENAVPHLRLTDEFTGTYHFD 420
Qy      418 EVFHSYSYASOSLDR-----LNNPLI--DQVLYLNRTONOSGSAQNKDLFSRGS 467
Db      421 TKPL--KLHSMQTNNSLGLPRLTLTEPTBGGHQHTLPAAVTRKGYHQITINNSYTEAT 478
Qy      468 PAGMSVQPKNMLPGPCYRQORVSKITDNNNSFTYTGASKYVLNGRESLINP--GTMA 525
Db      479 ---AIRP-----AQGVNTPYNNMFESNGGF-----LREIVPTADT 512
Qy      526 SHKDDKDFEPMGGMVIFGKESAGASNTALD---NMWITDEEIK--ATNP-----VAT 574
Db      513 QYNDDEPN-----GALRFTMDYHGHLITSSQLEKERTFNPQSKGRAPK 557
Qy      575 ERFGTVA-VNFQSS-----TDPATG--DVHMGALP-----GMWQDEYVYLQGEIWA 620
Db      558 QGFNQQAPLNLNTNNNGTLPSDPLGKSNMFMMTLNTYGPLTALNNTAPVPFNGQIWD 617
Qy      621 KLPHTD--GHFHSPLMGFGFLK-NPPQILIKNTVPANPPAEFSA-TKFASFITYOST 676
Db      618 KELDTDLKPRH---VTAPFVCCKNPPGQLFVKIAP---NLTDPMADSPOQPRITITSN 671
Qy      677 GQVSEIEMELOKENSKRMNPEVOYTSNYSKANSANVDFTVDNNGLYTEPRPIG 728
Db      672 FMMKGTTLFTFAKMRSSNMNMNFIQHTT-----TAENIGNYI-PTNIG 712

RESULT 7
coat protein VP1 - feline panleukopenia virus
N:Contains: coat protein VP2
C:Species: feline panleukopenia virus, FPLV
C>Date: 30-Jun-1987 #sequence_revision 30-Jun-1987 #text_change 09-Jul-2004
C:Accession: A03701
R:Carlson, J.; Rushlow, K.; Maxwell, I.; Maxwell, F.; Winston, S.; Hahn, W.
J. Virol. 55, 574-587, 1985
A>Title: Cloning and sequence of DNA encoding structural proteins of the autonomous par
A:Reference number: A03697; MUID:85265017; PMID:2391581
A:Accession: A03701
A:Molecule type: DNA
A:Residues: 1-727 <CAR>
A:Cross-references: UNIPROT:P04864; UNIPARC:UPI0000127D55; EMBL:M10824; NID:G333474; PI
C:Genetics:
A:Intons: 11/1
C:Superfamily: parvovirus coat protein
C:Keywords: coat protein
F:144-727/Product: coat protein VP2 #status predicted <VP2>

Query Match      11.1%; Score 444.5; DB 1; Length 727;
Best Local Similarity 23.7%; Pred. No. 4.6e-21;
Matches 187; Conservative 117; Mismatches 295; Indels 189; Gaps 35;

Qy      43 RGLVLPGYKYLGPFGNGLDKGEPVNADAALEHDKAVDQOLKAGDNFYLYRNHADAFOE 102
Db      10 RGLVPEGYKYLGPNSLSDGEPTNPDAAKEHDEAYAAALRGKKNFYLFSPADQRFID 69
Qy      103 RLQEDTSPFGNLGRAVFAQAKRVLEPLGLVEBGAKTAPGKKRPEVDSPOEDSSSGIGK 162
Db      70 QYKDADTWGKIGHYFFRAKKAIAVLTDPDHPSTSRPTK-PTKSKRPPIHFINIAK- 127
Qy      163 GQOPAKKRLNFGQTGSESV---DQPLGEPRA--TPAAVGPPTTMAAGGAPMAADNNE 216
Db      128 -----KKKAGAGVKKRDNQAPMSDGAQVDPGGQPAVNERATSGNGSGGG-----GG 176
Qy      217 GADGVNAGSNM-----CDSTWLGDVRVITTSRTMALPPTNNHLYKQI-----SS 262
Db      177 GSGGVGISTGTNNQTEPFKLENGWV--ETTAISSRLVHLMPESEYKVVVYVNMMDKTA 234
Qy      263 ASTGASNDNHYFGYSTPMGYPFDFNRFCHFSPPDMQRLINNMGFRPKRLNFKLFINIQV 322
Db      235 VKGNMALDDTHVQIVTFPMSLV DANAGWFPNPDMDLIVNTMSEILVLSFEQEIFVNLVK 294
```

Qy	323	EY	---	TNDGVT	TTANN	LTSTQV	FSDE	XYOL	PVYL	GAHQ	CL	-----	-PPFA	-----	368																																		
Db	225	TVSBA	TP	PKTVY	NNDL	TAS	LMV	LD	SNN	T	FF	PA	M	RSE	TLG	FY	WK	PI	LP	MY	354																												
Qy	369	--	DV	MI	PO	GY	LT	LN	NG	SO	AV	GR	S	-----	SFC	Y	C	L	E	-----	Y	PP	S	Q	M	L	T	G	NN	-	T	F	S	413															
Db	355	FQ	MD	TL	PI	SH	-----	TG	ST	G	P	T	NI	Y	H	G	T	D	D	D	D	D	D	D	D	D	D	D	D	D	D	D	D	D	408														
Qy	414	Y	TE	E	V	P	H	S	S	A	H	S	O	S	L	D	L	M	N	--	E	L	D	Q	Y	I	Y	N	R	O	N	O	S	G	A	N	O	N	D	L	E	F	R	S	P	A	G	M	471
Db	409	F	F	D	C	K	--	C	R	L	T	H	W	O	T	N	R	A	L	G	P	--	F	L	N	S	I	P	O	S	E	G	A	T	N	F	G	D	I	G	O	D	K	R	G	V	459		
Qy	472	S	-----	--	VQ	P	N	--	W	L	P	C	Y	R	O	R	V	S	K	-----	T	K	T	D	N	N	N	S	F	T	W	504																	

```

RESULT 8
VCPNAME      coat protein VP1 - mink enteritis virus (strain Abashiri)
N;Contains:   coat protein VP2
C;Species:    mink enteritis virus, MEV
C;Date:       30-Jun-1992 #sequence_revission 30-Jun-1992 #text_change 09-Jul-2004
C;Accession: B36350
R;Karatsumai, T.; Horiuchi, M.; Hama, E.; Yaguchi, K.; Ishiguro, N.; Goto, H.; Shingajima, S.; et al., 1997
J. Gen. Virol. 72, 867-875, 1991
A>Title: Construction and nucleotide sequence analysis of an infectious DNA clone of the mink enteritis virus strain Abashiri
A;Reference number: A38350; MUID:91202123; PMID:2016597
A;Accession: B36350
A:Molecule type: DNA
A;Residues: 1-722 <RAR>
A:CROSS-References: UNIPROT:P27437, UNIPARC:UPI0000174967, GB:D00765
C;Superfamily: parvovirus coat protein
C;Keywords: coat protein
F,i39-722/Product: coat protein VP2 #status predicted <VP2>

Query Match          11.0%; Score 439.5; DB 1; Length 722;
Beat Local Similarity 23.6%; Pred. No. 9.6e-21;
Matches 186; Conservative 117; Mismatches 295; Indels 189; Gaps 35;

QY  44 GLVLPEYKYTLGPENGLDGEFPVNADAAALSHDKAYPDQOLKAGNDPPLYRLYNADAFOER 103
     ||| |::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
Db   6 GLVPPPYKYTLGPFNSLDGSEPTNSSDAAKKHDEYAIVLRSGKNPIYLFSAPDRFIQQ 65

QY  104 LQEDTSSFGSNLGRAVFQAKKRVLEPLGLVEEGAKTAPEKKRPVEQSPOEPDSSSGIKTGG 163
     ::::||||:|:::||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
Db   66 TRDATDMGGIGHYFFRAKKAIAEVLRTDPNPSTSRSPETK--PTRKSKEPHIFINLAH-- 122

QY  164 QQPAAKRKLNGCGIDSSSV-----DQQLGEPFA---TTAAAVPTTMASGGCAPMADNNHG 217
     :::::::::::::||||:|::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
Db   123 ----KKKAGAGGVRYRDNLAFPSDGAVOPDGQPAVRNRBRATSGNGSGGGG-----GGG 172

QY  218 ADGVGNASGNWH-----CDSTWLGDGRVITTSFTWMLPTYNNHLVKYOI-----SSA 263

```

```

Db 173 SGGVOISITGCTFRNQNCFEFLKENGWV--ETANSRKLVLHMPSESEYKXVYVANNMDKTV 230
QY 264 STGASNDNHVFGYSGTSPWGYFDFNRFCHFSPPDMORLINNMGFRPKLNFLENQVKE 323
Db 221 KGNMALDOTHQOIVTPMSLVDNANAGVWFNPEDMOLIVNTWSEHLVSEFOELFNVLTXT 290
QY 324 VT---TNDGVTTIANLNTSTQVVFSDSEQLVYGVSAHQGL-----PPFA----- 368
Db 221 VSESGTQPTPKYVNNDLTASLWALDSSNNTMFTPAAMSETLGFYPMKPTTPTPKRYFY 350
QY 369 --DVFMIPQGYLTLLNNGSQAVGRS-----SFYCLE-YFPSQMLRTGNNP-TFSY 414
Db 351 QMDRLLIFSH-----TGTSGFPTNYHGTDEDDVDQFYTIENSVPVHLIRTGDEFPATGTF 404
QY 415 TFESEVPFHSYVAHQSLSRLMN--PLIQOVLVYLRTQNGSSAQOKDLFSRGSFAGMS 472
Db 405 FFDCKP--CRLTHWTQTNKALGLP-----FLNSPOSBGATNFGDGVQODKRGVY 455
QY 473 -----VQPKN-WLPGPCYRQORVSK-----TKTDNNNSNFTWTG 505
Db 456 QMGNTDYITEATIMRPAEVGYSAPFYSEASTQGFKEPIAARGGAGQTDENQA---DG 512
QY 506 ASKY---NLNRESTINRGT---AMAHKDEDEKFFPMSGWMIRCKEASAGSNTALD-N 557
Db 513 DPRVAFGRQHCQKTTTGETPERFTYIAHODT-----GRYPAGDMTQINFN 559
QY 558 VMTIDDEEIKATNPVATERFGTVAVNFOSSSTDPATGVDHAMGALPGWVMOBDDVYLOG 617
Db 560 LPVTDVNLVLPFDFIG---GKTGINY--TNIFNTYGPILTALNNVP-----PVPNQ 606
QY 618 IMAKIPIHTDGHFHPBLMGFGGLK-----NPPQIILLKNTVEPAN--PPAE 661
Db 607 IWDSEFDTD-----LKRLHVNAPFVCQNNQDGLFVYVAVNLNTNEYDDAS 653
QY 662 FSATKFAFSITQYSGQVSEIEMLOKENSKRANPEVOYTSNTYAKASANVDFTVDDNGLY 721
Db 654 ANMSR-----IYYSDFPMWKGLVFAKULRASHTWNPIDQMSIN-----VDNQFNY 699
QY 722 TEPAPIG 728
Db 700 L-PNNIG 705

```

```

RESULT 9
VCPVPF      coat protein VP1 - feline panleukopenia virus (strain 193)
N;Contains: coat protein VP2
C;Species: feline panleukopenia virus, FPLV
G;date: 31-Dec-1981 #sequence_revision 31-Dec-1991 #text_change 09-Jul-2004
J;Accession: B36608
R;Martyn, J.C.; Davidson, B.E.; Studdert, M.J.
U; Gen. Virol. 71, 2747-2753, 1990
A;Title: Nucleotide sequence of feline panleukopenia virus: comparison with canine parvovirus
A;Reference number: A36608; MUID:91073139; PMID:2174965
A;Accession: B36608
A:Molecule type: DNA
A;Residues: 1-727 <MAR>
A;Cross-references: UNIPROT:P24840; UNIPARC:UIP0000127D56; GB:X55115
C;Superfamily: parvovirus coat protein
C;Keywords: coat protein
F;144-727/Product: coat protein VP2 #status predicted <VP2>

Query Match          10.8%; Score 431.5; DB 1; Length 727;
Best Local Similarity    23.5%; Pred.No. 3.2e+20;
Matches 185; Conservative 117; Mismatches 297; Indels 189; Gaps 35;

Oy   43 RGLVLPGKYLYGPFGNGLDKGEPVNAADAAALEHDKAYQOOIKAGGNPLYRYNHADAEPQE 102
       |||||:::|||:::|||:::|||:::|||:::|||:::|||:::|||:::|||:::|||:::
Db   10 RGIVPPGYKYLYGFNSLDSGEPTNPSPDAALAEHDEAVYLRSGNNPLLTPSRAHQRFID  69
       :::::::::::::|||||::|||::|::|::|::|::|::|::|::|::|::|::|::|::|

Oy   103 RLQEDTFSGGNLGRAVPQAKKRVLIEPLGVGAKTAPDGKKRPVEQSPOEPPSSSGICKT  162
        : : : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
```

```

Db      70 QTKDAKMGCKIGHYFFRAKKAIAVLTTTPDHSTSPRTK-PTKRSKPPHIFINLAK- 127
QY      163 GOOPAKKRILNFGQTGDSSESV-----DPOPLGEPPA--TPAAVGPPTMASGGAPMADNNE 216
Db      128 -----KKKAGAGQVKRDNLAIPMSDGAIVQDGGQPAVRNERATSGSGGGG-----GG 176
QY      217 GADGVCNAGNNW-----CDSTWLGDRVITTTSTWMLPTYNHLYKOI-----SS 262
Db      177 GGGGVAISTGTFFNNQTEFFKLENGWV--ETTNASSSLVHLNMPESNRYKRVVNNNDKTA 234
QY      263 ASTGASNDHNYGCVSTPMGVPFDPNRPCHSPSPDMQRLINNNGWFRPKLNFLEFNIQVYK 322
Db      235 VAGNNALDIHQVITPMSLVDANANGWVFNPDMDLIYNTMSELHLVSEGEIIFNVVLK 294
QY      323 EYV-----TNDGVTTIANNLSTVQVFSDEXYQLPYVLSAHQGL-----PPPPA----- 368
Db      295 TVSEASATQPTPKVYNNDLTASLWALDLSNNTMPTPAAMSETLGFYPMKPTIPTPMRYX 354
QY      369 ---DVFMIPOYGLTLNNGSQAVGRS-----SYCYLE-YFPSSQMLRTGNNF-TFS 413
Db      355 FQWDRTLIPSH-----TGTSGTPTVYHGTDPDVQVFTTENSVPVHLRTGDEPATGT 408
QY      414 YTFEEVPFHSSVYAHQSGLRLNM--PLIDQVLYLNRTQNGSGSAQNKDLLFSRGSPPAGM 471
Db      409 FPFDCRP--CRLTHTWQTRALGLP-----FLNSLPSQEGATVFGDIGVQODKRGV 459
QY      472 S-----VOPKN-MLPGPCYRQORVSK-----TKTDNNNSNFTWT 504
Db      460 TQMGNTDVIYTEATIMRPAEVGYSAPYFSFASQTGFPTPIAAGRGAGQDENQAA---D 516
QY      505 GASKY---NLNGESLIINCT---AMASHKDEDEKFFPMSGWITGKSAGSANTALD- 556
Db      517 GEPRAVFGROHQKTTTGTGTPPERFTYIAHDT-----GRYPEGWIONIN 563
QY      557 NWMITDEEIKANPATERFGTVAVNFQSSSTDPATGVDHANGALPGMWQORDVYLOG 616
Db      564 NLPVIMDNVLLPFDPIG---GKTGINY--TNIFNTYGLTALANNP-----PYYPNG 610
QY      617 PIWAKIPIHTDGHFHPSPILMGFGFLK-----NPPQILIKNTVPVAN---PPA 660
Db      611 QIWDKEFTD-----LKPRLHVNAIPFVCCNCPGLFVYVAPNLITNEYDPDA 657
QY      661 EBSATFASFTIOYSTGQSVLEIEMELQKENSKRMPDEVQYTSNYAKSANVDTVUNNG 720
Db      658 SAAMSR---ITVYSDFPMWKGLVFPKAKLASHMTMPIOQMSIN-----VDNQEN 703
QY      721 YTEPRPIG 728
Db      704 YV-PNNIG 710

RESULT 10
VCPPVP
coat protein VP1 - porcine parvovirus
N:Contains: coat protein VP2
C:Species: porcine parvovirus
C>Date: 30-Sep-1991 #sequence_revision 30-Sep-1991 #text_change 09-Jul-2004
C/Accession: B33302
R/Panz, A.I.; Mancius, J.J.; Diaz-Aroca, E.; Casal, J.I.
J. Gen. Virol. 70, 2541-2553, 1989
A>Title: Porcine parvovirus: DNA sequence and genome organization.
A/Reference number: A33302; MUID:90010964; PMID:2794971
A/Molecule type: DNA
A/Residues: 1-723 <RAN>
A/Cross-references: UNIPROT:P18546; UNIPARC:UP10000174964; EMBL:DD0623
C:Genetics:
A/Intons: 10/1
C/Superfamily: parvovirus coat protein
C/Keywords: coat protein
F/145-723/Product: coat protein VP2 #status predicted <VP2>

Query Match      10.7%; Score 426; DB 1; Length 723;

```

```

Best Local Similarity 24.9%; Pred. No. 7,3e-20;
Matches 190; Conservative 109; Mismatches 308; Indels 156; Gaps 36;

QY      53 LGPFNGLDKGEVYNADAALAEHDKAYDQALAGDNPYLYRYNHADAEFOERLQEDTSFEG 112
Db      13 LGFNGSLDQGEPTNPDAAKHEIDAEYDKYISGKNPFTYFSADEKFIKETHAADYGG 72
QY      113 NIGRAVQAKKAVLEPLGLVEGAKTAPKPKPYVESPOEPDSSSIGK-----T 162
Db      73 KIGHYFFRAKRAFAPKL-----SETDSEPTTSQOEVRSPKHPQSKPPGKRPAPRHIFINL 129
QY      163 GOOPAKKRILNFGQTGDSSEVPDPOPLGEPPATP-AAVGPPTMASGGAPMADNNEGADGV 221
Db      130 AKKXAKGTINTNSNSNSENVQHNPTN--AAITELSTAGNESGGGGG-----GGRAGGV 183
QY      222 GNASGNW---CDSTWIGD--RVITTSSTWMLPTYNHLYKOI-----SSASTGAS-ND 270
Db      184 GYSTGSFNNQTEFGYLGEGLVRIITAHASRLIHLNMPHEHTYKRIHVLNESGSGAGQMVOD 243
QY      271 NHYFGYSTPWGYFDFNRFHCHSPSPDMQRLINNNGWFRPKLNFLEFNIQVYEVV---TN 327
Db      244 DAHTQMTVPMSLIDANAMGVNFPADWOLISNNMTEINLVSEQALFNVVLKTTESATS 303
QY      328 DGVTTIANNLSTVQVFSDEXYQLPYVLSAHQGLPPPADVFMIPQYGL----- 379
Db      304 PPKIYNNDLTASLWALDNTNLTPTTPAAPSSETLGFPMILPTKPTQYRYILSCIRNLN 363
QY      380 --TLNNGSQAVGRS-----FYCYLE-YFPSSQMLRTGNNF-TFSYFEEVPFHSSVYA 426
Db      364 PPTYGOSQPPNRLNTRNLHSDIMEFYTINAVPIHLFRGDERSTGIYHFDTKPL--KLT 421
QY      427 HGOSLDR-----LMNPLI--DOVLYLNRTQNGSGSAQNKDLLFSRGSPPAGMSVOPK 476
Db      422 HSWQNTKSLGLPRLKLTETTEGDHPGLPAANTKKGHQIINNSYTEAT---AIRP- 476
QY      477 NMLPGCYRQORVSKTKTDNNNSNFTWTGASKYNLNGRESIINP--GTMAASHKDEDEK 534
Db      477 -----AQGVNTPYNNFYSNGGPP-----LPIVPTADTYQNDDEPN- 514
QY      535 FPMGVMIFGKSAGSANTALD--NWMITDEEIK--ATNP-----VATERFGTVA-V 582
Db      515 -----GAIRFTMDYQHGHLTTSQSELEIRYFNPQSKGRAPKQOFNOQAPL 560
QY      583 NPOSS-----TPDATG--DVYAMGALP-----GNWQDROVYLOGPIWAKIPIHTD 627
Db      561 NLENTNNGTLTLPDDPIGKSNMHEFNTLNTYGPDLTALNTAPFENGQIWDKLDJDLKP 620
QY      628 HFPSPILMGFGFLK-NPPQILIKNTVPVANPPAEFSA-TKFASTIOYSTGQSVSEIEM 685
Db      621 RLH---VTAPFVCKNPPQGLPVKINP---NLTDDBFNADSPOQPRIITYISNPFMWKGTLP 674
QY      686 ELQKENSKRMPDEVQYTSNYAKSANVDTVUNNGLYTEPRPIG 728
Db      675 TAKMRSSNMNPIQCHT-----TAENIGKYI-PTNIG 706

RESULT 11
VCPVIM
coat protein VP1 - minute virus of mice (strain MW1)
C:Species: minute virus of mice, murine parvovirus
C>Date: 28-Dec-1987 #sequence_revision 28-Dec-1987 #text_change 09-Jul-2004
C/Accession: B23008; B29510
R/Schli, R.; McMaster, G.K.; Hirt, B.
Nucleic Acids Res. 13, 3617-3633, 1985
A>Title: DNA sequence comparison between two tissue-specific variants of the autonomous
A/Reference number: A23008; MUID:85242059; PMID:3855242
A/Molecule type: DNA
A/Residues: 1-718 <SAH>
A/Cross-references: UNIPROT:P07302; UNIPARC:UP10000174963; EMBL:X02481; NID:960918; PID:
J. Virol. 57, 656-669, 1986
A>Title: DNA sequence of the lymphotropic variant of minute virus of mice, MW(1), and

```

A:Reference number: A29510; MUID:86115415; PMID:3502703
A:Accession: B29510
A:Molecule type: DNA
A:Residues: 1-143, 'A', 145-718 <AST>
A:Cross-references: UNIPARC:UPI0000127D6C; EMBL:M12032
C:Superfamily: parvovirus coat protein
C:Keywords: coat protein

Query Match 10.6%; Score 424.5; DB 1; Length 718;
Best Local Similarity 23.7%; Pred. No. 9e-20;
Matches 184; Conservative 104; Mismatches 299; Indels 189; Gaps 31;

```
QY 45 LVLPFGYKILGPPNGLDKGEVNVNADAALEHDKAYDQQLKAGDNPYLRYNHADAEFQERL 104
D 1 MVRPGYKILGPGNSLDGEPNPSDAAKEHDEAVDQYKSGKNPYLYFSADQRFIDQ 60
QY 105 QEDTSFGCNLGRAVFOAKKRLVPLGLVEGAKTAPGKKRPVEOSPOEPDSSGIGKTG 163
D 61 KDAKDWGKGVGHYFFRTYRAFAPKLATDSE-----PG-----TSGVSRAGK 101
QY 164 -----QOPAKKRLNFG-----QTGDSSEVPDPQPLG-----BEPATPAVVGPTM 203
D 102 RTREPATVFINQAARAKKLTSSAAQSSQTMDSGTSGDGNVHSAARVERADGPG-- 159
QY 204 ASGGGAPMADNNEGADGVGNASGNWCHDS--TWLGD--RVITSTRTWALPTYNNHLKY 258
D 160 GSGGG-----GSGGGGAGVSTGSDYDQTHYRFLGDGVEITALATRLVHLNMPKSENYC 213
QY 259 QI-----SSASGASNDNHVFGYSTPFWGFFDFNRFCHSPRDMQGLINNMGFRKR 311
D 214 RIRVHNTTDTISVKNMAADHBEQIWFPMSLVDANAGVWLQPSDWQYICNTMSQLNVS 273
QY 312 LNFELFNIOVEKVTNTD-----GVTTIANLSTVQVESSEYQLPYVLGSAHQCLPFP 367
D 274 LDGIFNVVLTVEQDSGGAIKITNNNDLTACMMVANDSNMILPYTPAASMETLGYFP 333
QY 368 ADVPMIQQGY-----LTLNN-----GSOAVRSSFYCLEYFP--OMLRTG 407
D 334 WKPTIASPRYRYFCVDRDLSTVENOEGTIEHNVMGTPKGNMSOFFTIENTQOITLLRTG 393
QY 408 NNF--TFSYTFEEVPHFSSVAHSQSLDRLMN--PLIDQLYLYNRTQNGSQAOKDLIFS 464
D 394 DEFTIGTYFDPTNV--KLHTWQTNQQLGQPLSTP---PEADTDAGT---LTA 441
QY 465 RGSFAGNSVOPKMWL-----PPCYROORVSTKTDDNNNSN 500
D 442 QGSHHGATOMEVNVNVSARTRPAOVGRCPPHNDPEASRAGP--FAAPRVADVTQGVDR 500
QY 501 PTWTGASKYNLNGRESIINPGTAMASHKDEDKF-----FPMGCVMIFGESAGAS 551
D 501 ANGSVRSYSGKHGEMNAAHGPAPERVTWDETNGSGSDTRDGFQISAPLVVPPPLNGIL 560
QY 552 NTMLDNMWTDEEIKATNPVATERFGTVANFOS--SSTDPAQDVHAMALPGMWQD 609
D 561 TNA-----NPIGTKN---DIHFSNVFSYGPLAFSH-----P 590
QY 610 RDVYLQGPIMAK--IPHTDGHFSPPLMGFGKPNPQILIKNTP--VEANPAPAES 664
D 591 SPVYPQGIQIMKDLDEHKKRPLHTAPRV---CGNNAFGQMLVRLGRLDQVYRNQATL 647
QY 665 TKFASFTQYSTGVASVEIEMELQENSKRNPEVQTSNAYKASANDFTYDNNNG 719
D 648 SRIVTYGTFPMWKGKLTIRAKLRA---NTTNVPYQO-----VSVEDNG 686
```

RESULT 12

VCPCVP
coat protein VP1 - canine parvovirus (strain N)
N:Contains: coat protein VP2
C:Species: canine parvovirus, CPV
C:Date: 30-Jun-1989 #sequence_revision 30-Jun-1989 #text_change 09-Jul-2004
R:Reed, A.P.; Jones, E.V.; Miller, T.J.

U:Viro1. 62, 266-276, 1988
A:Title: Nucleotide sequence and genome organization of canine parvovirus.
A:Reference number: A29962; MUID:88062992; PMID:2824850
A:Accession: B29962
A:Molecule type: DNA
A:Residues: 1-748 <RE>
A:Cross-references: UNIPROT:P12930; UNIPARC:UPI0000127D7B; EMBL:M19296
C:Genetics:
A:Inserts: 26/3
C:Superfamily: parvovirus coat protein
C:Keywords: coat protein
P:165-748/Product: coat protein VP2 #status predicted <VP2>

Query Match 10.6%; Score 421.5; DB 1; Length 748;
Best Local Similarity 23.0%; Pred. No. 1.5e-19;
Matches 184; Conservative 114; Mismatches 287; Indels 215; Gaps 34;

```
QY 44 GLVLPGYKILGPPNGLDKGEVNVNADAALEHDKAYDQQLKAGDNPYLRYNHADAEFQER 103
D 32 GLVPPGYKILGPGNSLDGEPNPSDAAKEHDEAVYAAVLRSGKNPYLYFSADQRFIDQ 91
QY 104 LQEDTSFGCNLGRAVFOAKKRLVPLGLVEGAKTAPGKKRPVEOSPOEPDSSGIGKTG 163
D 92 TKDAKDWGKGIQHYFFRAKKAIAVLTDPHPBSTRTK--PTKRSKPPHIFILAK-- 148
QY 164 QOPAKKRLNFGQTDSEVP-----DQPLGPPA--TPAAVGPPTMASGGA PMADNNEG 217
D 149 -----KKKAGAGQVVRDNLAPWSDGAVQPDGQPAVRNBRATGSGNGSGGG-----GGG 198
QY 218 ADGVGNASGNW-----CDSTWLGDREVITSTRTWALPTYNNHLKYQI-----SSA 263
D 199 SGGAGISTGTFTNNQTEKFLENGWV--EITANSSRLVHLNMPSENYRNVVNNMDKAV 256
QY 264 STGASNDNHVFGYSTPFWGFFDFNRFCHSPRDMQGLINNMGFRKLNFLFNIOVEK 323
D 257 NGNMLDDIHAQIYTPMSLVDANAGVWFNCGDQILYNTMSELHVSFEOEIPNVVLKT 316
QY 324 VT---TNDGVTTIANLSTVQVESSEYQLPYVLGSAHQCL-----PPFPA----- 368
D 317 VSESATQPTKVYNNNDLTASLMLVALDSNNTMTPFPAAMRSBTLGFPWKPTIPFPMRYF 376
QY 369 --DVPMIQQGYLTLLNGSAVGRS-----SFCLE--YFPGQMLRTGNF--TFSY 414
D 377 QMDRTLLPSSH-----TGTSGTPNIYHGTDPDVOQFYTIENSVPVHLIRTGDEPATGTF 430
QY 415 TFEVPHFSSVAHSQSLDRLMN--PLIDQLYLYNRTQNGSQAOKDLFSRGSFAGMS 472
D 431 FFDCKP--CRLTHTWQTNRALGLRP-----FLNSLPQSGATNFGDI-----GV- 472
QY 473 VQPKMWLPGCYROORVSTKTDDNNNSN-----TWTGASKYNLNGRESIINP--GT 522
D 473 -----QDDKRGVATQMGNTNYITAEITMPAEVGYAPARYSFEASQGPFTK 519
QY 523 AMASHK-----DDEDKFPPMGCWMIFGESAGASTALD----- 556
D 520 PIAARGGAQTYENQAADGDPRY-----AFGQHQSKITTTETTERFTYIAHODTGR 572
QY 557 -----NVMITDEEIKATNPVATERFGTVANFOSSTDPAQDVHAMALPG 604
D 573 YPEBDWQININFLPVINDVNLFTDPIG--GKTGINY--TWIPNTYGLLTALNVP- 625
QY 605 MWQDRDVLVQGPIMAKIPHTDGHFSPPLMGFGK-----NPPQILIKN 651
D 626 -----FVYPNGQIWDKEPFD-----LKRPLHVNAPFVCONNCPQLFVKV 666
QY 652 TPVPAN--PPAESAKTFASFTQYSTGVASVEIEMELQENSKRNPEVQTSNAYKS 708
D 667 APNLTNEDPDASAMNR-----IYTSDFMWKGLVFKAKLRASHATWNPIDQMSIN----- 718
QY 709 ANVDFTVDNNGLYTEPRPIG 728
D 719 -----VDNQENYV--PSNIG 731
```


This Page Blank (uspto)

GenCore version 5.1.9
Copyright (c) 1993 - 2006 Bioceleleration Ltd.

OM protein - protein search, using sw model

Run on: June 12, 2006, 12:54:24 ; Search time 17 Seconds

(Without alignments)
547.376 Million cell updates/sec

Title: US-10-696-261-13

Perfect score: 3989 1 MADGYLEPMDLENDLSEGR.....NGLYTERPRIGTRYLRPL 736

Sequence:

Scoring table:

Searched: 64916 seqs, 12643201 residues

Total number of hits satisfying chosen parameters: 64916

Minimum DB seq length: 0
Maximum DB seq length: 200000000Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database:

Published Applications AA New:
1: /EMC_Celerra_SIDS3/prodata/2/pubppaa/US09_NEW_PUB.pap:*
2: /EMC_Celerra_SIDS3/prodata/2/pubppaa/US06_NEW_PUB.pap:*
3: /EMC_Celerra_SIDS3/prodata/2/pubppaa/US07_NEW_PUB.pap:*
4: /EMC_Celerra_SIDS3/prodata/2/pubppaa/US08_NEW_PUB.pap:*
5: /EMC_Celerra_SIDS3/prodata/2/pubppaa/PCT_NEW_PUB.pap:*
6: /EMC_Celerra_SIDS3/prodata/2/pubppaa/US10_NEW_PUB.pap:*
7: /EMC_Celerra_SIDS3/prodata/2/pubppaa/US11_NEW_PUB.pap:*
8: /EMC_Celerra_SIDS3/prodata/2/pubppaa/US01_NEW_PUB.pap:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	130	3.3	861	US-10-953-349-11809	Sequence 11809, A
2	130	3.3	887	US-10-953-349-11808	Sequence 11808, A
3	130	3.3	1004	US-10-953-349-11807	Sequence 11807, A
4	122.5	3.1	1629	US-10-471-571A-3542	Sequence 3542, A
5	119	3.0	667	US-10-505-928-393	Sequence 393, App
6	115	2.9	2271	US-10-471-571A-3948	Sequence 3948, App
7	106	2.7	2478	US-10-471-571A-2278	Sequence 2278, App
8	103.5	2.6	561	US-11-121-154-145	Sequence 145, App
9	102.5	2.6	684	US-11-293-697-4614	Sequence 4614, App
10	102.5	2.6	1464	US-10-985-570-2	Sequence 2, Appli
11	101.5	2.5	439	US-11-246-999-145	Sequence 145, App
12	101.5	2.5	1049	US-10-505-928-759	Sequence 759, App
13	101.5	2.5	1186	US-11-121-154-172	Sequence 172, App
14	99.5	2.5	375	US-10-953-349-23198	Sequence 23198, A
15	99.5	2.5	418	US-10-953-349-23197	Sequence 23197, A
16	99.5	2.5	513	US-10-505-928-765	Sequence 765, App
17	99.5	2.5	554	US-10-953-349-23196	Sequence 23196, A
18	99.5	2.5	1263	US-10-471-571A-5118	Sequence 5118, App
19	99	2.5	763	US-10-505-928-504	Sequence 504, App
20	98	2.5	374	US-10-953-349-33957	Sequence 33957, A
21	97	2.4	300	US-10-953-349-37333	Sequence 37333, A
22	97	2.4	388	US-10-953-349-30991	Sequence 30991, A
23	97	2.4	401	US-10-953-349-30990	Sequence 30990, A
24	97	2.4	421	US-10-953-349-10989	Sequence 10989, A
25	97	2.4	690	US-10-471-571A-1978	Sequence 1978, App

26	96.5	2.4	668	US-11-315-766-18	Sequence 18, Appl
27	95.5	2.4	1366	US-10-985-570-3	Sequence 3, Appli
28	95	2.4	346	US-10-953-349-28334	Sequence 28334, A
29	95	2.4	350	US-10-953-349-28333	Sequence 28333, A
30	95	2.4	362	US-10-511-937-4466	Sequence 4466, App
31	95	2.4	590	US-11-293-697-4216	Sequence 4216, App
32	95	2.4	1349	US-10-471-571A-3352	Sequence 3352, App
33	95	2.4	1466	US-10-985-570-1	Sequence 1, Appli
34	94.5	2.4	769	US-11-121-154-88	Sequence 88, Appl
35	94	2.4	964	US-11-134-228A-30	Sequence 30, Appl
36	94	2.4	4590	US-10-505-928-569	Sequence 569, App
37	93	2.3	562	US-11-313-356-7	Sequence 7, Appli
38	92	2.3	691	US-11-315-766-22	Sequence 22, Appl
39	92	2.3	748	US-11-293-697-3747	Sequence 3747, App
40	92	2.3	9535	US-10-471-571A-4496	Sequence 4496, App
41	91.5	2.3	372	US-11-293-697-4280	Sequence 4280, App
42	91.5	2.3	663	US-11-169-140-14	Sequence 14, Appl
43	91.5	2.3	1042	US-10-548-484-53	Sequence 53, Appl
44	90.5	2.3	1024	US-11-121-154-201	Sequence 201, App
45	90	2.3	381	US-11-321-421-8	Sequence 8, Appli

ALIGNMENTS

```
RESULT 1
US-10-953-349-11809
: Sequence 11809, Application US/10953349
: Publication No. US20060107345A1
: GENERAL INFORMATION:
: APPLICANT: ALEXANDROV, Nikolai et al.
: TITLE OF INVENTION: SEQUENCES-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTIDES
: FILE REFERENCE: 2750-1579US2
: CURRENT APPLICATION NUMBER: US/10/953,349
: NUMBER OF SEQ ID NOS: 40252
: SOFTWARE: PatentIn version 3.3
: SEQ ID NO 11809
: LENGTH: 861
: TYPE: PRT
: ORGANISM: Arabidopsis thaliana
US-10-953-349-11809

Query Match      3.3%; Score 130; DB 6; Length 861;
Best Local Similarity 21.0%; Pred. No. 0.066;
Matches 153; Conservative 69; Mismatches 289; Indels 216; Gaps 32;

OY 51 KYLGPFNGLDKGEFVNADDALEHDKAYDQQLKAGNPFYLRVNHADAFOERLOEDTSF 110
   |||
   |||
Db 17 KFLG-----NLPKKANSQDSYOKQ-EISRESVSREVLASSEKTDGAVDGTSK 64

OY 111 GGN-----LGRAVFOAKKRVLEPLGIVEGAKTAPGKKRPVEGSPQEP-----DSSSGIG 160
   |||
   |||
Db 65 TGSSKQDMKQKAYMPLTNSL-----EHPKKKSFRMSAHEDFLELDDIGAA 114

OY 161 KTCQOPAKKRLNFGQTSSEVPDPPLGEPPTPAAVGFTTMAAGGAPVADNNEGADG 220
   |||
   |||
Db 115 STPEVAEKNAFVEVESHLSI-----PIGKPLTPSAMSST-----SVISGDAASOG 163

OY 221 VGNAAGNMGHDSITWLGDRVITTTSTRTWALP-----TYNNHLKYQISSASTGASNDNHFGY 276
   |||
   |||
Db 164 TSNQS-----LETERNKFAFPIEAVQGSMASEPTSKFIQGTGKSSISSGK 210

OY 277 ST-----PMGYDFPFRFHGFPPRDMQRLINNMGFRPRRLNFKLNIQVKE 323
   |||
   |||
Db 211 PTSEKRIPLIEPKKPAAVP-----NISFSP-PATGLINNSG---ASADIKL-----EKT 258

OY 324 VTTNDGVTITANNLTSTVQVFS-----SEYQLPVLAGS-----AHQGL- 363
   |||
   |||
Db 259 SSTAFGVSEAMAKPTSEKTFNSAGSASTSAAPPLNLSIFSGAGANTVTPPSNGSLF 318
   |||
   |||
OY 364 --PPFADVFMIRFOYGYLTITLNGSQAVGRSSFYCLEIFPSOMLRGNNFTFSYFEEVFP 421
```

```

Db      319 SSSPFPRTISNP-----SDNSVGD-----MSTVQSFATNNSSTIFGKLP 361
Qy      422 HSSYHSSQSLDRLMNLIDQYLTLNRTQNOGSAQNKDLFGRSGPAG-MSVQPKWLP 480
Db      362 SMD-----SNSQSTASPLSSTSPFKFGQPAAPFSAFSAVSESS 399
Qy      481 GPCYRQQRVSKTYTDNNNSFTWTG-----ASKYNLNGRE----- 515
Db      400 GQISKETEY-KNATFENTSTFKFGMASADQSTGVSFGAKSAENKSRPGFVFGSSSVG 458
Qy      516 SIINPETAASHKDEDDKFFPMGCVMIFGKESAGASNTALDNVMTDEEIKATNPVARE 575
Db      459 STINPSTAAABES-----SGSLIFGVTSSTPGTETSKISAS--AATN-TGNS 506
Qy      576 RFGTVAVNFQSS-----STDPATGDVHAMGALPGWVWQDRDVLVLOGPIWAKI PHT 625
Db      507 VFGTSSPFTSSGSSWVGVSASTGSSVFGFNAVSASATSSQSAENRFGAGNAQTGMT 566
Qy      626 DG-----HFHPSPLMGFGGLKNPPQILKNTVPAN-PPAEFSATKFASTIQ 673
Db      567 GSGTTTSTQSIPIFGSSSPSAPSGLSG-----NSSLASNSPFGFSKSEPAVF-TS 617
Qy      674 YSTGQVS 680
Db      618 GSTPOL 624

```

RESULT 2

```

US-10-953-349-11808
; Sequence 11808, Application US/10953349
; Publication No. US20060107345A1
; GENERAL INFORMATION:
; APPLICANT: ALEXANDROV, Nikolai et al.
; TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTIDES
; FILE REFERENCE: 2750-1579PUS2
; CURRENT APPLICATION NUMBER: US/10/953,349
; CURRENT FILING DATE: 2004-09-30
; NUMBER OF SEQ ID NOS: 40252
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 11808
; LENGTH: 887
; TYPE: PRT
; ORGANISM: Arabidopsis thaliana
US-10-953-349-11808

```

Query Match 3.3%; Score 130; DB 6; Length 887;

Best Local Similarity 21.0%; Pred. No. 0.069; Mismatches 289; Indels 216; Gaps 32;

```

Db      51 KYLGPNGLDKGEPVNAADAALHDKAYDOOLKAGDNPYLRYNHADAEOERLOEDTSP 110
Qy      43 KFLG-----NLPEKKANSQSSYQKQ-EISRESVSREVLAQSEKGTGDAVDGTSK 90
Db      111 GGN-----LGRAVFOAKKRVLEPLGLVEGAKTAPCKRPVQSPQEP---DSSSGIG 160
Qy      91 TGSSKQODMKGKAVVMP/LTNSLE-----EHPKKKSPFMSAHEDFLELDDDDIGAA 140
Db      161 KTGOQPAKRLNGQGTDSSEVPDPOPLGEPATPAVGTPTMASGGGAPMADNNEGADG 220
Qy      141 STCEVAEKONAEVEKSHISI-----PIGKPLTPSEAMST-----SYSNGDSQG 169
Db      221 VGNASGNWCHDSTWLGDRVITTTSTRTWLP-----TYNNHLKYKQISSASTGASDNHFGY 276
Qy      190 TSNGS-----LETENKVFVAFIEAVQOSNMASEPTSKFIQGTSEKSSISSGK 236
Db      277 ST-----PWGYFDNRFCHGFSRDMQRLINNNWGRPRKPLNFKLPNIQVKE 323
Qy      237 PTSEBKRIPLIEPKKPAVFP-----NISFSP-PATGILNONG--ASADIKL-----EKT 284
Db      324 VTINDGVTTIANNLITSTVQVFS-----SEVQLPVLGS-----AHQGL- 363

```

```

Db      285 SSTARGVSEAMAKPTESKKTFSNSASGAESTSAAPTLNGSIFSAAGANTVTPPPNSGLT 344
Qy      364 --PPPADVFMIPQGYITLNNGSAVGRSSRYCLEYFQSOMLRGANNFTPTSEVEVP 421
Db      345 SSSPFPRTISNP-----SDNSVGD-----MSTVQSFATNNSSTIFGKLP 387
Qy      422 HSSYHSSQSLDRLMNLIDQYLTLNRTQNOGSAQNKDLFGRSGPAG-MSVQPKWLP 480
Db      388 SMD-----SNSQSTASPLSSTSPFKFGQPAAPFSAFSAVSESS 425
Qy      481 GPCYRQQRVSKTYTDNNNSFTWTG-----ASKYNLNGRE----- 515
Db      426 GQISKETEY-KNATFENTSTFKFGMASADQSTGVSFGAKSAENKSRPGFVFGSSSVG 484
Qy      516 SIINPETAASHKDEDDKFFPMGCVMIFGKESAGASNTALDNVMTDEEIKATNPVARE 575
Db      485 STINPSTAAABES-----SGSLIFGVTSSTPGTETSKISAS--AATN-TGNS 532
Qy      576 RFGTVAVNFQSS-----STDPATGDVHAMGALPGWVWQDRDVLVLOGPIWAKI PHT 625
Db      533 VFGTSSPFTSSGSSWVGVSASTGSSVFGFNAVSASATSSQSAENRFGAGNAQTGMT 592
Qy      626 DG-----HFHPSPLMGFGGLKNPPQILKNTVPAN-PPAEFSATKFASTIQ 673
Db      593 GSGTTTSTQSIPIFGSSSPSAPSGLSG-----NSSLASNSPFGFSKSEPAVF-TS 643
Qy      674 YSTGQVS 680
Db      644 GSTPOL 650

```

RESULT 3

```

US-10-953-349-11807
; Sequence 11807, Application US/10953349
; Publication No. US20060107345A1
; GENERAL INFORMATION:
; APPLICANT: ALEXANDROV, Nikolai et al.
; TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTIDES
; FILE REFERENCE: 2750-1579PUS2
; CURRENT APPLICATION NUMBER: US/10/953,349
; CURRENT FILING DATE: 2004-09-30
; NUMBER OF SEQ ID NOS: 40252
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 11807
; LENGTH: 1004
; TYPE: PRT
; ORGANISM: Arabidopsis thaliana
US-10-953-349-11807

```

Query Match 3.3%; Score 130; DB 6; Length 1004;

Best Local Similarity 21.0%; Pred. No. 0.083; Mismatches 289; Indels 216; Gaps 32;

```

Db      51 KYLGPNGLDKGEPVNAADAALHDKAYDOOLKAGDNPYLRYNHADAEOERLOEDTSP 110
Qy      160 KFLG-----NLPEKKANSQSSYQKQ-EISRESVSREVLAQSEKGTGDAVDGTSK 207
Db      111 GGN-----LGRAVFOAKKRVLEPLGLVEGAKTAPCKRPVQSPQEP---DSSSGIG 160
Qy      208 TGSSKQODMKGKAVVMP/LTNSLE-----EHPKKKSPFMSAHEDFLELDDDDIGAA 257
Db      161 KTGOQPAKRLNGQGTDSSEVPDPOPLGEPATPAVGTPTMASGGGAPMADNNEGADG 220
Qy      258 STCEVAEKONAEVEKSHISI-----PIGKPLTPSEAMST-----SYSNGDSQG 306
Db      221 VGNASGNWCHDSTWLGDRVITTTSTRTWLP-----TYNNHLKYKQISSASTGASDNHFGY 276
Qy      307 TSNGS-----LETENKVFVAFIEAVQOSNMASEPTSKFIQGTSEKSSISSGK 353
Db      277 ST-----PWGYFDNRFCHGFSRDMQRLINNNWGRPRKPLNFKLPNIQVKE 323
Qy      354 PTSEBKRIPLIEPKKPAVFP-----NISFSP-PATGILNONG--ASADIKL-----EKT 401

```

QY 324 VTTNDGVTTAANNITSTVOVFS-----SEYQPYVLGS-----AHQGL- 363
| : : : : :
Db 402 SSTAFGESEAMAKPTEBSKTFNSGASGASSTSAAPLNGSIFGAGANTVTPPSNGSLT 461
QY 364 --PPEPADVMIPQYGLITLNGSQAVGRSFCLEVFPPSQMLTGNFTFSYFEERYPF 421
| : : : : :
Db 462 SSBSFPPTISNIP-----SDNSVGD-----MPSTVQSPAAITHNSSIFGLKLP 504
QY 422 HSSAYASQSLDRMLNPLIDYLYLNRTQNSGSAONKDLIFSRSAG-MSVQPKWLP 480
| : : : : :
Db 505 SND-----SNSQTSASPLSSTSPFKGQAPAPSAVSESS 542
QY 481 GPCYRQORVSKTKTDNNNSNFTWTG-----ASKYNLNGRE----- 515
| : : : : :
Db 543 GQISKEDEV-KNAFTFGNTSTFKFGKMASADQSTGVGAKSAENKSRPGVFGSSVVG 601
QY 516 SIINPGTAMASHODEKFPFMGCVMTFGKESAGASTALDNWMTDEBEIKATNPVATE 575
| : : : : :
Db 602 STINPSTAAAPES-----SGSLIFGVTSSTPGETSKISASS--AAIN-TGNS 649
QY 576 RFGTAVNFOSS-----STDPATGDVHAMGALPGWVQDRDVLQGPWAKIPIHT 625
| : : : : :
Db 650 VFGTSSPFTSSSGSMWGVASSTGSSVFGFNVAVSASATSSQASNRFGAGAOIGNT 709
QY 626 DG-----HFHDSPLMGFGGLKNPPQILIKNTVPVAN-PPAEFSATKEASFTQ 673
| : : : : :
Db 710 GSGTTSTQSI PQFGSSPAPSPGLSG-----NSSLASNGSPFGFSKSEPAVF-TS 760
QY 674 YSTGQVS 680
| : : : : :
Db 761 GSTPQLS 767

RESULT 4
US-10-471-571A-3542
; Sequence 3542, Application US/10471571A
; Publication No. US20060115490A1
; GENERAL INFORMATION:
; APPLICANT: CHIRON SPA
; TITLE OF INVENTION: STAPHYLOCOCCUS AUREUS PROTEINS AND NUCLEIC ACIDS
; FILE REFERENCE: P026927NO
; CURRENT APPLICATION NUMBER: US/10/471.571A
; PRIOR FILING DATE: 2003-09-12
; PRIOR APPLICATION NUMBER: GB-0107661.1
; PRIOR FILING DATE: 2001-03-27
; NUMBER OF SEQ ID NOS: 5642
; SOFTWARE: SeqMan99, version 1.03
; SEQ ID NO 3542
; LENGTH: 1629
; TYPE: PRP
; ORGANISM: Staphylococcus aureus
; FEATURE:
; NAME/KEY: MISC_FEATURE
; LOCATION: (1)..(1629)
; OTHER INFORMATION: hypothetical protein
US-10-471-571A-3542

Query Match 3.1%; Score 122.5; DB 6; Length 1629;
Best Local Similarity 20.1%; Pred. No. 0.58; Indels 251; Gaps 41;
Matches 158; Conservative 100; Mismatches 278;
QY 45 LVLPGYKYLGPFGNLDKGEFVNADAAALHEDKAYDQOLKAGDNPYLRYNHADAEPQERL 104
| : : : : :
Db 36 ILIGSLMYLG-----TQGEAEAAENNIENPTTLK--DN---VQSKVKIEVY 78
QY 105 QEOTSFGAGNIGRAVFAQAKKVL--EPGLVBEAGAKTPGKKRPVQSPQEDSSSGTGKT 162
| : : : : :
Db 79 NKDTAPQGVAKSEVSNKDTIEHEPVGAKEDISKEDTPEKADVAVQPKKS--VTNH 136
QY 163 GQOPAKKRL-----NFGOTGDSSEVPD-----PQDLGPRPA 193
| : : : : :
Db 137 AETPKYRKAKASVDEGSDITRDSKNVVESTPIITIGCKEHEFGYGVNDIQKKPDLGLGSEV 196

QY 194 TPAAVGPTTMAAGGAPMAD-----NNEGADG-----VGNASGN 227
| : : : : :
Db 197 TRFVNGNESNGLIGALQKNKIDSKDFNFYKRVANNHQSSNTGADGGLFSGKNAB-E 255
QY 228 WHCSTLGDVRVITTSRTALPT---YNNHLYQIASASGASNDNHYPGYSTPWWGFPD 284
| : : : : :
Db 256 YLTNGGILGDKGLVNS-GGFKIDIGYIYTSMDTEKQAGG-----YROY---GAF- 303
QY 285 FNRPHCHSPRDWQRLINNMGFRPKRLNFKLFIQVKEVTTNDGVTTIANNITSTVOVF 344
| : : : : :
Db 304 -----VANDSG-----NSQWGENIDKSKTN--FLNAYDNSTNT----- 336
QY 345 SDSEYQ-----LPYVLGSAHQG-----LPPEPADVMIPQYGLYL 381
| : : : : :
Db 337 SDGKFGQORLNDVILTVASTGKKRAEYAGKTWETSITDGLSKQOAVNPLITSQRMGL 396
QY 382 NNGSQAVGRSFCLEVFPPQMLT---GNFTFS-----YFEFVPPHSSYAH 427
| : : : : :
Db 397 NQGINANG-----WMRTDKGSEFTTPEAPKTITTELEKKVEEIPFKK--- 439
QY 428 SOSLDRLNPLIDQYLYLNRTQNSG-----SAQN--KDLFSRGP-AGMSVQPKN 477
| : : : : :
Db 440 -----ERKFNPDLAGTEKAYR-EGQKGEKTTTPTLKNPLTGVLIISKGEPEKITXPDIN 494
QY 478 WLP-----GPCYRQORVSKTKTDNNNSNFTWTGASKYNLNGRESIINP--GTAMASHK 528
| : : : : :
Db 495 ELTEYGEETIAPGHRDEFDPLPT-----GEKEVPKPGIKNPETDGVVPRPV 543
QY 529 DDEKFPFMGCVMTFGKESAGASTALDNWMTDEBEIKATNPVA--TERGCTAVANQS 586
| : : : : :
Db 544 DSVTKYGPVKGDSIVEKE-----IPEKEREKFNPDLAGTEKVTREGQKGEK 591
QY 587 SST-----DPATGDVHAMGALPGWVQDRDVLQGPWAKIPIHTDGHFAP-----S 632
| : : : : :
Db 592 TITPTLKNPLTGEIISKESKEITKDPINELTEYGETTTPGHRD-EFPKPLPTGKE 650
QY 633 PLMGFGGLKNPPQILIKNTVPVANPPAEFSATKFASTIOYSGQVSEIEMELQKENS 692
| : : : : :
Db 651 EVPGKPGIKNPETDVVR-----PVVD-SVTXYGPV-----KGDSIVEKE-EIPFEKE 696
QY 693 KRNPEV 699
| : : : : :
Db 697 RKNPDL 703

RESULT 5
US-10-505-928-393
; Sequence 393, Application US/10505928
; Publication No. US20060088532A1
; GENERAL INFORMATION:
; APPLICANT: Ludwig Institute for Cancer Research et al.
; TITLE OF INVENTION: LYMPHATIC ENDOTHELIAL GENES
; FILE REFERENCE: 28967/39178
; CURRENT APPLICATION NUMBER: US/10/505.928
; PRIOR FILING DATE: 2004-08-27
; PRIOR APPLICATION NUMBER: US 60/363,019
; PRIOR FILING DATE: 2002-03-07
; NUMBER OF SEQ ID NOS: 866
; SOFTWARE: Patentin 3.2
; SEQ ID NO 393
; LENGTH: 667
; TYPE: PRP
; ORGANISM: Homo sapiens
US-10-505-928-393

Query Match 3.0%; Score 119; DB 6; Length 667;
Best Local Similarity 20.9%; Pred. No. 0.3; Indels 192; Gaps 33;
Matches 118; Conservative 75; Mismatches 179;
QY 191 PPAATPAVGPPTMASG--GGAPMADNNEGADGVGNASGNHCOSTWLGADVITTSRTTWA 248
| : : : : :
Db 28 PVSNGKNGPTSLASGHFTGSNVEDR-----SSGSM-----GNGGHPSPSKNYG 72


```
QY 359 HOGCLPPPADVPMIPOYGYTLNNGQAVGRSSFCLEYPSCMLRTGNFTSYFEE 418
; Sequence 145, Application US/11211154
; Publication No. US20060105914A1
; GENERAL INFORMATION:
; APPLICANT: TAYLOR, LARRY EDMUND
; APPLICANT: WEINER, RONALD M.
; APPLICANT: HUTCHESON, STEVEN WAYNE
; APPLICANT: EKBORG, NATHAN A.
; APPLICANT: HOWARD, MICHAEL
; TITLE OF INVENTION: PLANT WALL DEGRADATIVE COMPOUNDS AND SYSTEMS
; FILE REFERENCE: 108172-00121
; CURRENT APPLICATION NUMBER: US/11/121,154
; PRIOR FILING DATE: 2005-05-04
; PRIOR APPLICATION NUMBER: 60/567,971
; PRIOR FILING DATE: 2004-05-04
; NUMBER OF SEQ ID NOS: 214
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 145
; LENGTH: 561
; TYPE: PRT
; ORGANISM: Microbulbifer degradans
US-11-121-154-145

Query Match 2.6%; Score 103.5; DB 7; Length 561;
Best Local Similarity 19.6%; Pred. No. 3.2;
Matches 70; Conservative 57; Mismatches 157; Indels 73; Gaps 14;

QY 280 WGYDFRFRHCFPRBDMQRLINNNGW--FRPKRLNFKLEIQKVEYTTNGVTIANNL 337
; Sequence 145, Application US/11211154
; Publication No. US20060105914A1
; GENERAL INFORMATION:
; APPLICANT: TAYLOR, LARRY EDMUND
; APPLICANT: WEINER, RONALD M.
; APPLICANT: HUTCHESON, STEVEN WAYNE
; APPLICANT: EKBORG, NATHAN A.
; APPLICANT: HOWARD, MICHAEL
; TITLE OF INVENTION: PLANT WALL DEGRADATIVE COMPOUNDS AND SYSTEMS
; FILE REFERENCE: 108172-00121
; CURRENT APPLICATION NUMBER: US/11/121,154
; PRIOR FILING DATE: 2005-05-04
; PRIOR APPLICATION NUMBER: 60/567,971
; PRIOR FILING DATE: 2004-05-04
; NUMBER OF SEQ ID NOS: 214
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 145
; LENGTH: 561
; TYPE: PRT
; ORGANISM: Microbulbifer degradans
US-11-121-154-145

QY 201 WG-----QDMQVMLNNASEVAQADSLSTMTFSVHMVQVYNN--LSTVENYV 245
; Sequence 145, Application US/11211154
; Publication No. US20060105914A1
; GENERAL INFORMATION:
; APPLICANT: TAYLOR, LARRY EDMUND
; APPLICANT: WEINER, RONALD M.
; APPLICANT: HUTCHESON, STEVEN WAYNE
; APPLICANT: EKBORG, NATHAN A.
; APPLICANT: HOWARD, MICHAEL
; TITLE OF INVENTION: PLANT WALL DEGRADATIVE COMPOUNDS AND SYSTEMS
; FILE REFERENCE: 108172-00121
; CURRENT APPLICATION NUMBER: US/11/121,154
; PRIOR FILING DATE: 2005-05-04
; PRIOR APPLICATION NUMBER: 60/567,971
; PRIOR FILING DATE: 2004-05-04
; NUMBER OF SEQ ID NOS: 214
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 145
; LENGTH: 561
; TYPE: PRT
; ORGANISM: Microbulbifer degradans
US-11-121-154-145

QY 338 TSTVQVFSDESEYQLPYVL---GSAHQGLPPPADVPMIPOY--GYLTLN---NGS----- 385
; Sequence 145, Application US/11211154
; Publication No. US20060105914A1
; GENERAL INFORMATION:
; APPLICANT: TAYLOR, LARRY EDMUND
; APPLICANT: WEINER, RONALD M.
; APPLICANT: HUTCHESON, STEVEN WAYNE
; APPLICANT: EKBORG, NATHAN A.
; APPLICANT: HOWARD, MICHAEL
; TITLE OF INVENTION: PLANT WALL DEGRADATIVE COMPOUNDS AND SYSTEMS
; FILE REFERENCE: 108172-00121
; CURRENT APPLICATION NUMBER: US/11/121,154
; PRIOR FILING DATE: 2005-05-04
; PRIOR APPLICATION NUMBER: 60/567,971
; PRIOR FILING DATE: 2004-05-04
; NUMBER OF SEQ ID NOS: 214
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 145
; LENGTH: 561
; TYPE: PRT
; ORGANISM: Microbulbifer degradans
US-11-121-154-145

QY 246 STFL-----SSHNLPLVGEFGADHQEVEDAILSAEYGYGYLQMSNGSGSCCGT 300
; Sequence 145, Application US/11211154
; Publication No. US20060105914A1
; GENERAL INFORMATION:
; APPLICANT: TAYLOR, LARRY EDMUND
; APPLICANT: WEINER, RONALD M.
; APPLICANT: HUTCHESON, STEVEN WAYNE
; APPLICANT: EKBORG, NATHAN A.
; APPLICANT: HOWARD, MICHAEL
; TITLE OF INVENTION: PLANT WALL DEGRADATIVE COMPOUNDS AND SYSTEMS
; FILE REFERENCE: 108172-00121
; CURRENT APPLICATION NUMBER: US/11/121,154
; PRIOR FILING DATE: 2005-05-04
; PRIOR APPLICATION NUMBER: 60/567,971
; PRIOR FILING DATE: 2004-05-04
; NUMBER OF SEQ ID NOS: 214
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 145
; LENGTH: 561
; TYPE: PRT
; ORGANISM: Microbulbifer degradans
US-11-121-154-145

QY 386 -QAVGRSSFCLEYFPCOMLRTGNFTSYFEEVPHSYAHSQSLDRLNPLIDQYLY 444
; Sequence 145, Application US/11211154
; Publication No. US20060105914A1
; GENERAL INFORMATION:
; APPLICANT: TAYLOR, LARRY EDMUND
; APPLICANT: WEINER, RONALD M.
; APPLICANT: HUTCHESON, STEVEN WAYNE
; APPLICANT: EKBORG, NATHAN A.
; APPLICANT: HOWARD, MICHAEL
; TITLE OF INVENTION: PLANT WALL DEGRADATIVE COMPOUNDS AND SYSTEMS
; FILE REFERENCE: 108172-00121
; CURRENT APPLICATION NUMBER: US/11/121,154
; PRIOR FILING DATE: 2005-05-04
; PRIOR APPLICATION NUMBER: 60/567,971
; PRIOR FILING DATE: 2004-05-04
; NUMBER OF SEQ ID NOS: 214
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 145
; LENGTH: 561
; TYPE: PRT
; ORGANISM: Microbulbifer degradans
US-11-121-154-145

QY 301 LDITNNFNVNSLITSMGRLLNGTNGIKATSVIASYVSGSSSSSSSSSS----- 348
; Sequence 145, Application US/11211154
; Publication No. US20060105914A1
; GENERAL INFORMATION:
; APPLICANT: TAYLOR, LARRY EDMUND
; APPLICANT: WEINER, RONALD M.
; APPLICANT: HUTCHESON, STEVEN WAYNE
; APPLICANT: EKBORG, NATHAN A.
; APPLICANT: HOWARD, MICHAEL
; TITLE OF INVENTION: PLANT WALL DEGRADATIVE COMPOUNDS AND SYSTEMS
; FILE REFERENCE: 108172-00121
; CURRENT APPLICATION NUMBER: US/11/121,154
; PRIOR FILING DATE: 2005-05-04
; PRIOR APPLICATION NUMBER: 60/567,971
; PRIOR FILING DATE: 2004-05-04
; NUMBER OF SEQ ID NOS: 214
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 145
; LENGTH: 561
; TYPE: PRT
; ORGANISM: Microbulbifer degradans
US-11-121-154-145

QY 445 YLNTQNGSQAOKDILLFSRGPAGMSVQPKMLPGPCY-----RQGRVS 490
; Sequence 145, Application US/11211154
; Publication No. US20060105914A1
; GENERAL INFORMATION:
; APPLICANT: TAYLOR, LARRY EDMUND
; APPLICANT: WEINER, RONALD M.
; APPLICANT: HUTCHESON, STEVEN WAYNE
; APPLICANT: EKBORG, NATHAN A.
; APPLICANT: HOWARD, MICHAEL
; TITLE OF INVENTION: PLANT WALL DEGRADATIVE COMPOUNDS AND SYSTEMS
; FILE REFERENCE: 108172-00121
; CURRENT APPLICATION NUMBER: US/11/121,154
; PRIOR FILING DATE: 2005-05-04
; PRIOR APPLICATION NUMBER: 60/567,971
; PRIOR FILING DATE: 2004-05-04
; NUMBER OF SEQ ID NOS: 214
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 145
; LENGTH: 561
; TYPE: PRT
; ORGANISM: Microbulbifer degradans
US-11-121-154-145

QY 349 --SSTSSSSSSSTSSSSSSSSSSSSG--AQCCNMY--GSVYPLCNNQASGKMGWQOQCTIG 403
; Sequence 145, Application US/11211154
; Publication No. US20060105914A1
; GENERAL INFORMATION:
; APPLICANT: TAYLOR, LARRY EDMUND
; APPLICANT: WEINER, RONALD M.
; APPLICANT: HUTCHESON, STEVEN WAYNE
; APPLICANT: EKBORG, NATHAN A.
; APPLICANT: HOWARD, MICHAEL
; TITLE OF INVENTION: PLANT WALL DEGRADATIVE COMPOUNDS AND SYSTEMS
; FILE REFERENCE: 108172-00121
; CURRENT APPLICATION NUMBER: US/11/121,154
; PRIOR FILING DATE: 2005-05-04
; PRIOR APPLICATION NUMBER: 60/567,971
; PRIOR FILING DATE: 2004-05-04
; NUMBER OF SEQ ID NOS: 214
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 145
; LENGTH: 561
; TYPE: PRT
; ORGANISM: Microbulbifer degradans
US-11-121-154-145

QY 491 KTKTDNNNSNFTMGASKYKYNLNGRESLIPGTAMASHKDDKFFPMGVMIPGKESAGA 550
; Sequence 145, Application US/11211154
; Publication No. US20060105914A1
; GENERAL INFORMATION:
; APPLICANT: TAYLOR, LARRY EDMUND
; APPLICANT: WEINER, RONALD M.
; APPLICANT: HUTCHESON, STEVEN WAYNE
; APPLICANT: EKBORG, NATHAN A.
; APPLICANT: HOWARD, MICHAEL
; TITLE OF INVENTION: PLANT WALL DEGRADATIVE COMPOUNDS AND SYSTEMS
; FILE REFERENCE: 108172-00121
; CURRENT APPLICATION NUMBER: US/11/121,154
; PRIOR FILING DATE: 2005-05-04
; PRIOR APPLICATION NUMBER: 60/567,971
; PRIOR FILING DATE: 2004-05-04
; NUMBER OF SEQ ID NOS: 214
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 145
; LENGTH: 561
; TYPE: PRT
; ORGANISM: Microbulbifer degradans
US-11-121-154-145

QY 404 RTTESSQSGNGGVIGSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSG 463
; Sequence 145, Application US/11211154
; Publication No. US20060105914A1
; GENERAL INFORMATION:
; APPLICANT: TAYLOR, LARRY EDMUND
; APPLICANT: WEINER, RONALD M.
; APPLICANT: HUTCHESON, STEVEN WAYNE
; APPLICANT: EKBORG, NATHAN A.
; APPLICANT: HOWARD, MICHAEL
; TITLE OF INVENTION: PLANT WALL DEGRADATIVE COMPOUNDS AND SYSTEMS
; FILE REFERENCE: 108172-00121
; CURRENT APPLICATION NUMBER: US/11/121,154
; PRIOR FILING DATE: 2005-05-04
; PRIOR APPLICATION NUMBER: 60/567,971
; PRIOR FILING DATE: 2004-05-04
; NUMBER OF SEQ ID NOS: 214
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 145
; LENGTH: 561
; TYPE: PRT
; ORGANISM: Microbulbifer degradans
US-11-121-154-145

QY 551 SNTLDNVMTDEDEIKATNPVATERGCTVAVNFQSSSTDPATGDIAMGALPGMNV 607
; Sequence 145, Application US/11211154
; Publication No. US20060105914A1
; GENERAL INFORMATION:
; APPLICANT: TAYLOR, LARRY EDMUND
; APPLICANT: WEINER, RONALD M.
; APPLICANT: HUTCHESON, STEVEN WAYNE
; APPLICANT: EKBORG, NATHAN A.
; APPLICANT: HOWARD, MICHAEL
; TITLE OF INVENTION: PLANT WALL DEGRADATIVE COMPOUNDS AND SYSTEMS
; FILE REFERENCE: 108172-00121
; CURRENT APPLICATION NUMBER: US/11/121,154
; PRIOR FILING DATE: 2005-05-04
; PRIOR APPLICATION NUMBER: 60/567,971
; PRIOR FILING DATE: 2004-05-04
; NUMBER OF SEQ ID NOS: 214
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 145
; LENGTH: 561
; TYPE: PRT
; ORGANISM: Microbulbifer degradans
US-11-121-154-145

QY 464 SGATCEHI-ITNSWNSGFOGAVRITNNGSSAINGVQVSW-----YSDGTITGSVW 513
; Sequence 145, Application US/11211154
; Publication No. US20060105914A1
; GENERAL INFORMATION:
; APPLICANT: TAYLOR, LARRY EDMUND
; APPLICANT: WEINER, RONALD M.
; APPLICANT: HUTCHESON, STEVEN WAYNE
; APPLICANT: EKBORG, NATHAN A.
; APPLICANT: HOWARD, MICHAEL
; TITLE OF INVENTION: PLANT WALL DEGRADATIVE COMPOUNDS AND SYSTEMS
; FILE REFERENCE: 108172-00121
; CURRENT APPLICATION NUMBER: US/11/121,154
; PRIOR FILING DATE: 2005-05-04
; PRIOR APPLICATION NUMBER: 60/567,971
; PRIOR FILING DATE: 2004-05-04
; NUMBER OF SEQ ID NOS: 214
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 145
; LENGTH: 561
; TYPE: PRT
; ORGANISM: Microbulbifer degradans
US-11-121-154-145

RESULT 9
US-11-293-697-4614
```

```
; Sequence 4614, Application US/11293697
; Publication No. US20060105376A1
; GENERAL INFORMATION:
; APPLICANT: HELIX RESEARCH INSTITUTE
; TITLE OF INVENTION: Novel full length cdna
; FILE REFERENCE: H1-A0106
; CURRENT APPLICATION NUMBER: US/11/293,697
; CURRENT FILING DATE: 2005-12-05
; PRIOR APPLICATION NUMBER: US/10/108,260
; PRIOR FILING DATE: 2002-03-28
; NUMBER OF SEQ ID NOS: 5458
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 4614
; LENGTH: 684
; TYPE: PRT
; ORGANISM: Homo sapiens
US-11-293-697-4614

Query Match 2.6%; Score 102.5; DB 7; Length 684;
Best Local Similarity 25.9%; Pred. No. 5;
Matches 68; Conservative 28; Mismatches 90; Indels 77; Gaps 15;

QY 26 KCGAPKPKRANQKDDGRGLVLPGYKYLGPENGLDKG-EPVNAADAALHDKAYDQ-QL 83
; Sequence 4614, Application US/11293697
; Publication No. US20060105376A1
; GENERAL INFORMATION:
; APPLICANT: HELIX RESEARCH INSTITUTE
; TITLE OF INVENTION: Novel full length cdna
; FILE REFERENCE: H1-A0106
; CURRENT APPLICATION NUMBER: US/11/293,697
; CURRENT FILING DATE: 2005-12-05
; PRIOR APPLICATION NUMBER: US/10/108,260
; PRIOR FILING DATE: 2002-03-28
; NUMBER OF SEQ ID NOS: 5458
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 4614
; LENGTH: 684
; TYPE: PRT
; ORGANISM: Homo sapiens
US-11-293-697-4614

QY 156 KPKDPKPKVKKLY-----HQYIPDQKAEKPPWDASAYARLLQQQLFLQLQI 205
; Sequence 4614, Application US/11293697
; Publication No. US20060105376A1
; GENERAL INFORMATION:
; APPLICANT: HELIX RESEARCH INSTITUTE
; TITLE OF INVENTION: Novel full length cdna
; FILE REFERENCE: H1-A0106
; CURRENT APPLICATION NUMBER: US/11/293,697
; CURRENT FILING DATE: 2005-12-05
; PRIOR APPLICATION NUMBER: US/10/108,260
; PRIOR FILING DATE: 2002-03-28
; NUMBER OF SEQ ID NOS: 5458
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 4614
; LENGTH: 684
; TYPE: PRT
; ORGANISM: Homo sapiens
US-11-293-697-4614

QY 84 KAGDNPYLRYNHADAERQERLQEDTSPGNGLGRAVFOAKKRVLEPLGLVEGAKTAPGKK 143
; Sequence 4614, Application US/11293697
; Publication No. US20060105376A1
; GENERAL INFORMATION:
; APPLICANT: HELIX RESEARCH INSTITUTE
; TITLE OF INVENTION: Novel full length cdna
; FILE REFERENCE: H1-A0106
; CURRENT APPLICATION NUMBER: US/11/293,697
; CURRENT FILING DATE: 2005-12-05
; PRIOR APPLICATION NUMBER: US/10/108,260
; PRIOR FILING DATE: 2002-03-28
; NUMBER OF SEQ ID NOS: 5458
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 4614
; LENGTH: 684
; TYPE: PRT
; ORGANISM: Homo sapiens
US-11-293-697-4614

QY 206 L-----SQQQQQQHQHRSYLG-----MQNQ-----LX 228
; Sequence 4614, Application US/11293697
; Publication No. US20060105376A1
; GENERAL INFORMATION:
; APPLICANT: HELIX RESEARCH INSTITUTE
; TITLE OF INVENTION: Novel full length cdna
; FILE REFERENCE: H1-A0106
; CURRENT APPLICATION NUMBER: US/11/293,697
; CURRENT FILING DATE: 2005-12-05
; PRIOR APPLICATION NUMBER: US/10/108,260
; PRIOR FILING DATE: 2002-03-28
; NUMBER OF SEQ ID NOS: 5458
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 4614
; LENGTH: 684
; TYPE: PRT
; ORGANISM: Homo sapiens
US-11-293-697-4614

QY 144 RPEVQSQFOEPDSSS-GIGKTGQOPAKKRLNFGQGDSESVDDPOPLGEPPATPAVGPPT 202
; Sequence 4614, Application US/11293697
; Publication No. US20060105376A1
; GENERAL INFORMATION:
; APPLICANT: HELIX RESEARCH INSTITUTE
; TITLE OF INVENTION: Novel full length cdna
; FILE REFERENCE: H1-A0106
; CURRENT APPLICATION NUMBER: US/11/293,697
; CURRENT FILING DATE: 2005-12-05
; PRIOR APPLICATION NUMBER: US/10/108,260
; PRIOR FILING DATE: 2002-03-28
; NUMBER OF SEQ ID NOS: 5458
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 4614
; LENGTH: 684
; TYPE: PRT
; ORGANISM: Homo sapiens
US-11-293-697-4614

QY 229 EPNEQWVRNPSSSTPLSNTPLSFKNSFS-GQTGVSSFKEGPPY---PNLDDLKVSRLR 284
; Sequence 4614, Application US/11293697
; Publication No. US20060105376A1
; GENERAL INFORMATION:
; APPLICANT: HELIX RESEARCH INSTITUTE
; TITLE OF INVENTION: Novel full length cdna
; FILE REFERENCE: H1-A0106
; CURRENT APPLICATION NUMBER: US/11/293,697
; CURRENT FILING DATE: 2005-12-05
; PRIOR APPLICATION NUMBER: US/10/108,260
; PRIOR FILING DATE: 2002-03-28
; NUMBER OF SEQ ID NOS: 5458
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 4614
; LENGTH: 684
; TYPE: PRT
; ORGANISM: Homo sapiens
US-11-293-697-4614

QY 203 M-----ASGGAPADN-NEGADGVGNASGMHCHDSTLGGPVTITRTMALPTYN 253
; Sequence 4614, Application US/11293697
; Publication No. US20060105376A1
; GENERAL INFORMATION:
; APPLICANT: HELIX RESEARCH INSTITUTE
; TITLE OF INVENTION: Novel full length cdna
; FILE REFERENCE: H1-A0106
; CURRENT APPLICATION NUMBER: US/11/293,697
; CURRENT FILING DATE: 2005-12-05
; PRIOR APPLICATION NUMBER: US/10/108,260
; PRIOR FILING DATE: 2002-03-28
; NUMBER OF SEQ ID NOS: 5458
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 4614
; LENGTH: 684
; TYPE: PRT
; ORGANISM: Homo sapiens
US-11-293-697-4614

QY 285 QQLRIRGLPVSGYTTALMDRLRPQDCSGNPNF-----GD-ITVT---FVTP 331
; Sequence 4614, Application US/11293697
; Publication No. US20060105376A1
; GENERAL INFORMATION:
; APPLICANT: HELIX RESEARCH INSTITUTE
; TITLE OF INVENTION: Novel full length cdna
; FILE REFERENCE: H1-A0106
; CURRENT APPLICATION NUMBER: US/11/293,697
; CURRENT FILING DATE: 2005-12-05
; PRIOR APPLICATION NUMBER: US/10/108,260
; PRIOR FILING DATE: 2002-03-28
; NUMBER OF SEQ ID NOS: 5458
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 4614
; LENGTH: 684
; TYPE: PRT
; ORGANISM: Homo sapiens
US-11-293-697-4614

QY 254 NHLVKQISASTGA-SNDNHYG 275
; Sequence 4614, Application US/11293697
; Publication No. US20060105376A1
; GENERAL INFORMATION:
; APPLICANT: HELIX RESEARCH INSTITUTE
; TITLE OF INVENTION: Novel full length cdna
; FILE REFERENCE: H1-A0106
; CURRENT APPLICATION NUMBER: US/11/293,697
; CURRENT FILING DATE: 2005-12-05
; PRIOR APPLICATION NUMBER: US/10/108,260
; PRIOR FILING DATE: 2002-03-28
; NUMBER OF SEQ ID NOS: 5458
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 4614
; LENGTH: 684
; TYPE: PRT
; ORGANISM: Homo sapiens
US-11-293-697-4614

QY 332 NTLPNYQSSSTALSNGFYHFG 354
; Sequence 4614, Application US/11293697
; Publication No. US20060105376A1
; GENERAL INFORMATION:
; APPLICANT: HELIX RESEARCH INSTITUTE
; TITLE OF INVENTION: Novel full length cdna
; FILE REFERENCE: H1-A0106
; CURRENT APPLICATION NUMBER: US/11/293,697
; CURRENT FILING DATE: 2005-12-05
; PRIOR APPLICATION NUMBER: US/10/108,260
; PRIOR FILING DATE: 2002-03-28
; NUMBER OF SEQ ID NOS: 5458
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 4614
; LENGTH: 684
; TYPE: PRT
; ORGANISM: Homo sapiens
US-11-293-697-4614

RESULT 10
US-10-985-570-2

Query Match 2.6%; Score 102.5; DB 6; Length 1464;
Best Local Similarity 23.4%; Pred. No. 15;
Matches 57; Conservative 6; Mismatches 62; Indels 119; Gaps 9;

QY 26 KCGAPKPKRANQKDDGRGLV---LPGYKYLGPENGID-----KGEVNAADA 70
; Sequence 4614, Application US/11293697
; Publication No. US20060105376A1
; GENERAL INFORMATION:
; APPLICANT: HELIX RESEARCH INSTITUTE
; TITLE OF INVENTION: Novel full length cdna
; FILE REFERENCE: H1-A0106
; CURRENT APPLICATION NUMBER: US/11/293,697
; CURRENT FILING DATE: 2005-12-05
; PRIOR APPLICATION NUMBER: US/10/108,260
; PRIOR FILING DATE: 2002-03-28
; NUMBER OF SEQ ID NOS: 5458
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 4614
; LENGTH: 684
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-985-570-2

QY 237 KPGRPGRGPRGPQ-GARGLPGTAGLPQMKGNHGRFSGLDGAKGDPAGPAGPGEPSGEN 295
; Sequence 4614, Application US/11293697
; Publication No. US20060105376A1
; GENERAL INFORMATION:
; APPLICANT: HELIX RESEARCH INSTITUTE
; TITLE OF INVENTION: Novel full length cdna
; FILE REFERENCE: H1-A0106
; CURRENT APPLICATION NUMBER: US/11/293,697
; CURRENT FILING DATE: 2005-12-05
; PRIOR APPLICATION NUMBER: US/10/108,260
; PRIOR FILING DATE: 2002-03-28
; NUMBER OF SEQ ID NOS: 5458
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 4614
; LENGTH: 684
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-985-570-2

QY 71 AALEHDAKAYOQLKAGDNPYLRYNHADAERQERLQEDTSPGNGLGRAVFOAKKRVLEPLG 130
; Sequence 4614, Application US/11293697
; Publication No. US20060105376A1
; GENERAL INFORMATION:
; APPLICANT: HELIX RESEARCH INSTITUTE
; TITLE OF INVENTION: Novel full length cdna
; FILE REFERENCE: H1-A0106
; CURRENT APPLICATION NUMBER: US/11/293,697
; CURRENT FILING DATE: 2005-12-05
; PRIOR APPLICATION NUMBER: US/10/108,260
; PRIOR FILING DATE: 2002-03-28
; NUMBER OF SEQ ID NOS: 5458
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 4614
; LENGTH: 684
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-985-570-2
```

Db 296 GA-----PGMG-----PRG 305
QY 131 LV-EEGAKTAPGKKRPEVQSPQEPDSSSGTGTGQOPAKRLINFGOTGSESVDPDQPLG 189
Db 306 LPERGRPGAPG-----PAGARGNDGATG--AAGPFGPTGPA 340
QY 190 EPPATPAV-----GPTMASGGA PMADNNEGADGVG 222
Db 341 GPPGFGAAGAKGAEAPQGRGSEPGVGRGEPGPPGAGAPAGNPGADGQPGAKGAN 400
QY 223 NASG 226
Db 401 GAPG 404
RESULT 11
US-11-246-999-145
; Sequence 145, Application US/11246999
; Publication No. US20060099622A1
; GENERAL INFORMATION:
; APPLICANT: NI et al.
; TITLE OF INVENTION: 12 Human Secreted Proteins
; FILE REFERENCE: PF489P2
; CURRENT APPLICATION NUMBER: US/11/246,999
; PRIOR FILING DATE: 2005-10-11
; PRIOR APPLICATION NUMBER: US/09/984,130
; PRIOR FILING DATE: 2001-10-29
; PRIOR APPLICATION NUMBER: 60/243,792
; PRIOR FILING DATE: 2000-10-30
; PRIOR APPLICATION NUMBER: 09/836,353
; PRIOR FILING DATE: 2001-04-18
; PRIOR APPLICATION NUMBER: 60/198,407
; PRIOR FILING DATE: 2000-04-19
; PRIOR APPLICATION NUMBER: PCT/US99/25031
; PRIOR FILING DATE: 1999-10-27
; PRIOR APPLICATION NUMBER: 60/105,971
; PRIOR FILING DATE: 1998-10-28
; NUMBER OF SEQ ID NOS: 149
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 145
; LENGTH: 439
; TYPE: PRT
; ORGANISM: Homo sapiens
US-11-246-999-145
Query Match 2.5%; Score 101.5; DB 7; Length 439;
Best Local Similarity 20.9%; Pred. No. 3.2;
Matches 71; Conservative 49; Mismatches 90; Indels 129; Gaps 20;
QY 293 SPEDMQRLLNNMGFRPKR-----LNFKLFNIQVKEVTNNGV-----TTIAN 335
Db 53 NPGGFTFLDNHWGLDQRYRLIHYSEDELIRLSNITVHD-----EGVYCYYSSTPFRS 108
QY 336 NLTS-----TVQVFSDSSEYQLPYVLGSAHQGLPPPADVFWIPOYGYLTLANGS 385
Db 109 KMTTVEVLAPSKRPVLOVSDTEGRV--TLSCYTQCKPQ-----PQITWL-LDNGI 157
QY 386 QAVG-----RSSFYCLEYPPSQMLRT-----GNNFTFSYTFEEVPPH 422
Db 158 QLEGDTRHKLLEADQKWTTLTTLVLAYGPNSTATCLVHHKALGGCKLTPQFEDVA-- 215
QY 423 SSVAHQSLDRKLNPLIDVYL-VYLNRTQNOGSAQKDLLFSRGSAPAGSVOPKMLPG 481
Db 216 RYANNTPTVSTTLE--VDIVSEVQPTVTTAESDLNSNDFP----- 256
QY 482 PCYRQGRVSKTKTDNNNSNFT-----WTGASKYNL-NGRESIINPTAMASHKDEDKF 534
Db 257 PSIFPQ-----HNSGATSVAAAGELSGTSAHHIEGTETALN-CTV-----TEELF 300
QY 535 -----FPMGVMIFGKESAGASNTALDNVMTDEEETIKA 568
Db 301 RTEASFSEENVTLIS-----IVTFEQDVKS 325

RESULT 12
US-10-505-928-759
; Sequence 759, Application US/10505928
; Publication No. US20060088532A1
; GENERAL INFORMATION:
; APPLICANT: Ludwig Institute for Cancer Research et al.
; TITLE OF INVENTION: LYMPHATIC ENDOTHELIAL GENES
; FILE REFERENCE: 28967/39178
; CURRENT APPLICATION NUMBER: US/10/505,928
; PRIOR FILING DATE: 2004-08-27
; PRIOR APPLICATION NUMBER: US 60/363,019
; PRIOR FILING DATE: 2002-03-07
; NUMBER OF SEQ ID NOS: 866
; SOFTWARE: PatentIn 3.2
; SEQ ID NO 759
; LENGTH: 1049
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-505-928-759
Query Match 2.5%; Score 101.5; DB 6; Length 1049;
Best Local Similarity 18.2%; Pred. No. 11;
Matches 122; Conservative 79; Mismatches 179; Indels 291; Gaps 35;
QY 122 KKRVLRLPLGVEGA-KTAPGKKRPVQSPQEPDS-----SSGIGKTGOOPAK-KRLN 172
Db 82 KANTSOP-GVLQGAAYLLCPWGAASPTQCTPIEDFSKGRLESLSSEGEVEVEYKSLQ 140
QY 173 -FGQT-----GDSESVDP-----QPLGEPATPAVGPPTMASGG-----GAPMD 213
Db 141 WRGATVRAHAGSSILACAPLYSMWTEKEPLSDP-----VGTCTSLTDNTRLILEVAPCRS 194
QY 214 NNEGADGVGNASNNHCDSST-----WIGDKVITSTRTMALPTN-----N 254
Db 195 DPSMAAGOGYCOGFGFAEFTKGRVVLGGPGSYFWOQO--LISATOEQIAESYPPYLYN 252
QY 255 HLYKOISSASTGASNNHYFGYSTPMGYFDENRFHCFSPRMDQRLINNMGFRPYRLNF 314
Db 253 LVGGQLQIROASSIYDSTLIGISVAVG-----EFGSD----- 285
QY 315 KLFNIQKEVTNNDGVTTI-ANNLSTVQVFSDSSEYQLPYVLGSAHQGLPPPADVFWI 373
Db 286 -----TEDFVAGVPKGNLT----- 299
QY 374 POYGYLTLANGSAVGRSSFYCLEYPPSQMLRTGNNFTSYTFEEVPHSSYA-----H 427
Db 300 --YGYVTIILNGSD-----IRS-----LYNFSGEQMASYFGYVAATDVN 336
QY 428 SOSLDRLM--NPLIDQYLYLNRTQNOGSAQKDLLFSRGSAPAGSVOPKMLPGPCYR 485
Db 337 GGDGLDLVLGAPFL-----MDRTPPGRPOEVGRVYVYLQ-HRAGIEPPT----- 379
QY 486 OGRVSKTKTDNNNSNFTWTGASKYKYNLNGRESIINPTAMASHKD-DEKFFPMGVMIFG 544
Db 380 -----TLTIGHDEF--GRF-----GSSILPLDLDLDODGVDVAIGAFAFG 417
QY 545 KESAGASNTALDNVMTDEBEIKATNPATERGTYAANVFQSSSTPPATGVDVAMAGALPG 604
Db 418 GEF-----QGGVVF-----PFGGPGGLSGSKPS 440
QY 605 MWQODRDVYLQGIWAKIPHT-----DGHFSPPLMGFGGLKNPPOLILK 650
Db 441 QVLIQ-----PLMA-ASHTPDFGSALRGGRDLDGNGYFDLIVGSFGVDK--AVVYR 488
QY 651 NPEV-----PA-----EPASATKFAFTIQTSGQVSEI 683
Db 489 GRPIVASASLITFPAMFNPBERSCSLGEPVACINLSFCINASGKHVADSIGFTVELQL 548
QY 684 EWELOKENSKR 694
Db 549 DMQOKQGVRR 559


```
RESULT 13
US-11-121-154-172
; Sequence 172, Application US/11121154
; Publication No. US20060105914A1
; GENERAL INFORMATION:
; APPLICANT: TAYLOR, LARRY EDMOND
; APPLICANT: WEINER, RONALD M.
; APPLICANT: HUTCHESON, STEVEN WAYNE
; APPLICANT: EKBORG, NATHAN A.
; APPLICANT: HOWARD, MICHAEL
; TITLE OF INVENTION: PLANT WALL DEGRADATIVE COMPOUNDS AND SYSTEMS
; FILE REFERENCE: 108172-00121
; CURRENT APPLICATION NUMBER: US/11/121,154
; PRIOR FILING DATE: 2005-05-04
; PRIOR APPLICATION NUMBER: 60/567,971
; NUMBER OF SEQ ID NOS: 214
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO: 172
; LENGTH: 1186
; TYPE: PRT
; ORGANISM: Microbulbifer degradans
US-11-121-154-172

Query Match
Best Local Similarity 19.8%; Pred. No. 13;
Matches 85; Conservative 50; Mismatches 139; Indels 155; Gaps 21;

QY 349 YQAPYVIGSAMHQLGPPF-----PADVFMTPQVGYLT-----LNNG 384
D 211 YRTQVVGGEASVPTPTHKMGVSFLGEMKIGDPAGA-----GYTPDPTARISKG 264
QY 385 SOAVGRSSFYCLEFFPSQMLRTGNNFTFSY--TFEEVPHSSYASQSLDRIMNPLIDQY 442
D 265 VRIMG-----IPNLGAGNQGYTSVDPDFSEVDDGIAVANSEYAN-----LEAY 309
QY 443 L-YLLNRT--QNOGSGAQNKDLLFSRGSFAGMSVQPKNMLPGCYRQQRYSKTKTDNNN 498
D 310 LKSYSDGTATVQMOGSLPVMQATFVHGSP-----YVFPKAYRGNMVLRTKADCG 360
QY 499 SNFTWTGASKYNLNGRESIINPGTAMASHKDEDEKFFPMGVMFGKESAASVTALDNV 558
D 361 EKGTF-----YNNNSLGI--WTSVANKNKD-----FLITGEGEVFNNIETDTI 403
QY 559 MITD-----BEEKATNPVA-----TERFGTVAV 582
D 404 TLTTAANEFTLTLLPTAGAGTPSSTVIOAFEDSARAV--VAKDIOYSVDRTNNMVTYTH 461
QY 583 NFOGSSSTDPAATGDVHANGALPGWVQDRVYLQ-----PIWAKIPHTDGHFPPSLMG 636
D 462 TYKNESNTP--VQTLAGLLPYMHMKYSDTALSGYKTRSARGMVQFAHIDFSYTIPIYVG 517
QY 637 GFGLKANPPOLLKNTVPANPPAFESATKRFASFTIOYSTQGVSEVLEMELOKNSKRN 696
D 518 -----VL-----PYLPSSVGDFFDSVLAGLVQAFVA-----GSPENMN 550
QY 697 PEVOYTSNY 705
D 551 P---HTDTY 556

RESULT 14
US-10-953-349-23198
; Sequence 23198, Application US/10953349
; Publication No. US20060107345A1
; GENERAL INFORMATION:
; APPLICANT: ALEXANDROV, Nickolai et al.
; TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTIDES
; FILE REFERENCE: 2750-1579PUS2
; CURRENT APPLICATION NUMBER: US/10/953,349
; CURRENT FILING DATE: 2004-09-30
```

```
; NUMBER OF SEQ ID NOS: 40252
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 23198
; LENGTH: 375
; TYPE: PRT
; ORGANISM: Glycine max
US-10-953-349-23198

Query Match
Best Local Similarity 19.8%; Pred. No. 3.6;
Matches 63; Conservative 21; Mismatches 97; Indels 137; Gaps 14;

QY 5 GYLPDWLEDNLSGIREMWDLKPGAPKPKANQKODDGRGLVLPQYKYLQ--PENGLDKG 62
D 131 GYPP-----QQPDGYSTWMDOSTAP--HQSTHGG-----GYYYSSQOPQOPQPNPG 175
QY 63 EPVNAADAALAEHDK-----AYDQOLKAG--DNPTLYRNHADAERQERLOEDTSFGNLR 116
D 219 GPAPPADGSAYNYSQPPSSGYNQSQAQGYAODSTNAYN-----AQSGSGYG----- 264
QY 117 AVFOAKRVLLEPLGLVEGAKTAPGKKRPVEQSPQEPDSSSGIGTKGOQPAKK--RLNFG 174
D 265 -----GPTVYDQOQGYG--SSSNPAQEGHTANYA 291
QY 175 QTGDSSEVPDPQPL-----GEPATPAVAGPTTMA 204

RESULT 15
US-10-953-349-23197
; Sequence 23197, Application US/10953349
; Publication No. US20060107345A1
; GENERAL INFORMATION:
; APPLICANT: ALEXANDROV, Nickolai et al.
; TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTIDES
; FILE REFERENCE: 2750-1579PUS2
; CURRENT APPLICATION NUMBER: US/10/953,349
; NUMBER OF SEQ ID NOS: 40252
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 23197
; LENGTH: 418
; TYPE: PRT
; ORGANISM: Glycine max
US-10-953-349-23197

Query Match
Best Local Similarity 19.8%; Pred. No. 4.2;
Matches 63; Conservative 21; Mismatches 97; Indels 137; Gaps 14;

QY 5 GYLPDWLEDNLSGIREMWDLKPGAPKPKANQKODDGRGLVLPQYKYLQ--PENGLDKG 62
D 174 GYPP-----QQPDGYSTWMDOSTAP--HQSTHGG-----GYYYSSQOPQOPQPNPG 218
QY 63 EPVNAADAALAEHDK-----AYDQOLKAG--DNPTLYRNHADAERQERLOEDTSFGNLR 116
D 219 GPAPPADGSAYNYSQPPSSGYNQSQAQGYAODSTNAYN-----AQSGSGYG----- 264
QY 117 AVFOAKRVLLEPLGLVEGAKTAPGKKRPVEQSPQEPDSSSGIGTKGOQPAKK--RLNFG 174
D 265 -----GPTVYDQOQGYG--SSSNPAQEGHTANYA 291
QY 175 QTGDSSEVPDPQPL-----GEPATPAVAGPTTMA 204
```

Db	292	VQGD\$AQA\$P\$AQPTVTAQXGYPTNQLPSSNTANYPPQGTTPQPGYGVPT\$QAAVGNQ\$OP	351
Qy	205	SGGGA\$PMADNNE-----GADGVGNASGNMHCDSTWLGDRVITSTRWALPTYNNHL	256
Db	352	GYGGA\$VQPGYGPQTYGAPQGGQPGYQ-----ALPSYSNS	387
Qy	257	Y---KOISSASTGASNDN	271
Db	388	YGAGYTQTPAYTGDNGN	405

Search completed: June 12, 2006, 12:57:44
Job time : 19 secs

GenCore version 5.1.9
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: June 12, 2006, 12:43:59 ; Search time 305 Seconds
(without alignments)
2232.169 Million cell updates/sec

Title: US-10-696-261-13

Perfect score: 3989
Sequence: 1 MADGYLPDWLEDNLSGIR.....NNGLYTEPRDITGRYLTRPL 736

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 2849598 seqs, 925015592 residues

Total number of hits satisfying chosen parameters: 2849598

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Uniprot_7.2.*

1: uniprot_sprot.*

2: uniprot_trembl.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	3989	100.0	736	Q9WB88_VVIRU	Q9WB88 adeno-ssoc
2	3968	99.5	736	Q6JC08_VVIRU	Q6JC08 adeno-ssoc
3	3963	99.3	736	Q6J137_VVIRU	Q6J137 adeno-ssoc
4	3883.5	97.4	737	Q6JC13_VVIRU	Q6JC13 adeno-ssoc
5	3873	97.1	736	Q6JC12_VVIRU	Q6JC12 adeno-ssoc
6	3853	96.6	736	Q6JC10_VVIRU	Q6JC10 adeno-ssoc
7	3511	88.0	736	Q6J139_VVIRU	Q6J139 adeno-ssoc
8	3494	87.6	736	Q6J131_VVIRU	Q6J131 adeno-ssoc
9	3473	87.1	736	Q6J08X3_VVIRU	Q6J08X3 non-human P
10	3469.5	87.0	735	Q6J008_VVIRU	Q6J008 adeno-ssoc
11	3456.5	86.7	735	Q6J877_VVIRU	Q6J877 adeno-ssoc
12	3456.5	86.7	737	Q6JC52_VVIRU	Q6JC52 adeno-ssoc
13	3455.5	86.6	737	Q6JC58_VVIRU	Q6JC58 adeno-ssoc
14	3453.5	86.6	735	Q6JC28_VVIRU	Q6JC28 adeno-ssoc
15	3453	86.6	738	Q6JC19_VVIRU	Q6JC19 adeno-ssoc
16	3451.5	86.5	735	Q6JC34_VVIRU	Q6JC34 adeno-ssoc
17	3450.5	86.5	737	Q6JC51_VVIRU	Q6JC51 adeno-ssoc
18	3449.5	86.5	735	Q6JB21_VVIRU	Q6JB21 adeno-ssoc
19	3449.5	86.5	735	Q6JC42_VVIRU	Q6JC42 adeno-ssoc
20	3449	86.5	738	Q5Y9B4_VVIRU	Q5Y9B4 adeno-ssoc
21	3449	86.5	738	Q6JC37_VVIRU	Q6JC37 adeno-ssoc
22	3448	86.4	738	Q808W5_VVIRU	Q808W5 non-human P
23	3446.5	86.4	735	Q6JC17_VVIRU	Q6JC17 adeno-ssoc
24	3445	86.4	738	Q6JC60_VVIRU	Q6JC60 adeno-ssoc
25	3443.5	86.3	737	Q6JC46_VVIRU	Q6JC46 adeno-ssoc
26	3442.5	86.3	735	Q6JC36_VVIRU	Q6JC36 adeno-ssoc
27	3442.5	86.3	737	Q8J0G0_VVIRU	Q8J0G0 adeno-ssoc
28	3442	86.3	738	Q6JC14_VVIRU	Q6JC14 adeno-ssoc
29	3442	86.3	738	Q6JC56_VVIRU	Q6JC56 adeno-ssoc
30	3441.5	86.3	735	Q6JC44_VVIRU	Q6JC44 adeno-ssoc
31	3440	86.2	738	Q6JC15_VVIRU	Q6JC15 adeno-ssoc

32	3440	86.2	738	2	Q6JC16_VVIRU	Q6JC16 adeno-ssoc
33	3440	86.2	738	2	Q6JC62_VVIRU	Q6JC62 adeno-ssoc
34	3439	86.2	738	2	Q6JB22_VVIRU	Q6JB22 adeno-ssoc
35	3438	86.2	734	2	Q6JC02_VVIRU	Q6JC02 adeno-ssoc
36	3438	86.2	738	2	Q808X3_VVIRU	Q808X3 non-human P
37	3437	86.2	738	2	Q6JC47_VVIRU	Q6JC47 adeno-ssoc
38	3437	86.2	738	2	Q6JC61_VVIRU	Q6JC61 adeno-ssoc
39	3436	86.1	734	2	Q6JC04_VVIRU	Q6JC04 adeno-ssoc
40	3434.5	86.1	735	2	Q6JB26_VVIRU	Q6JB26 adeno-ssoc
41	3434	86.1	728	2	Q808X4_VVIRU	Q808X4 non-human P
42	3434	86.1	738	2	Q6J8X4_VVIRU	Q6J8X4 adeno-ssoc
43	3433	86.1	738	2	Q6JC49_VVIRU	Q6JC49 adeno-ssoc
44	3433	86.1	738	2	Q6JC53_VVIRU	Q6JC53 adeno-ssoc
45	3432	86.0	738	2	Q6JC54_VVIRU	Q6JC54 adeno-ssoc

ALIGNMENTS

RESULT 1	Q9WB88_VVIRU	PRELIMINARY; PRT; 736 AA.
ID	Q9WB88_VVIRU	
AC	Q9WB88;	
DT	01-NOV-1999, integrated into UniprotKB/TREMBL.	
DT	01-NOV-1999, sequence version 1.	
DT	07-FEB-2006, entry version 18.	
DE	Capsid protein.	
OS	Adeno-associated virus - 1.	
OC	Viruses; ssDNA viruses; Parvoviridae; Parvovirinae; Dependovirus.	
OX	NCBI_TaxID=85106;	
RN	[1]	
RP	NCLEOTIDE SEQUENCE.	
RX	MEDLINE=99214338; Pubmed=10196295;	
RA	Xiao W., Chirmule N., Berta S.C., McCullough B., Gao G., Wilson J.M.;	
RT	"Gene therapy vectors based on adeno-associated virus type 1."	
RL	J. Virol. 73:3994-4003(1999).	
CC	-----	
CC	Copyrighted by the Uniprot Consortium, see http://www.uniprot.org/terms	
CC	Distributed under the Creative Commons Attribution-NonDerivs License	
CC	-----	
DR	EMBL; AF063497; AAD27757.1; -; Genomic_DNA.	
DR	SMR; Q9WB88; 217-736.	
DR	GO; GO:0019028; C:Viral capsid; IEA.	
DR	GO; GO:0005198; F:structural molecule activity; IEA.	
DR	InterPro; IPR001403; Parvo_coat.	
DR	Pfam; PF00740; Parvo_coat; 1.	
SO	SEQUENCE 736 AA; 81375 MW; CFABFB9B5DC0595 CRC64;	
Query Match	100.0%; Score 3989; DB 2; Length 736;	
Best Local Similarity	100.0%; Pred. No. 4e-251;	
Matches	736; Conservative 0; Mismatches 0; Indels 0; Gaps 0;	
Qy	1	MADGYLPDWLEDNLSGIREWMDLKGAPKPKANQKODDGRGLVPGYKYLGPFGD 60
Db	1	MADGYLPDWLEDNLSGIREWMDLKGAPKPKANQKODDGRGLVPGYKYLGPFGD 60
Qy	61	KGEFVNAADAALAEHDKAYDQOLKAGDNPYLRYNHADAERQERLQEDTSPFGNIGRAVFQ 120
Db	61	KGEFVNAADAALAEHDKAYDQOLKAGDNPYLRYNHADAERQERLQEDTSPFGNIGRAVFQ 120
Qy	121	AKKRVLEPLGLVEGATTAAGKKRPVQSQEQEPDSSSGIGTKGQPPAKKRLNFQGTGDS 180
Db	121	AKKRVLEPLGLVEGATTAAGKKRPVQSQEQEPDSSSGIGTKGQPPAKKRLNFQGTGDS 180
Qy	181	SVDPPOPGLGEPPTPAVGPPTMASGGGAPMADNNEGADGVGNAGNMHCSTVLGDRV 240
Db	181	SVDPPOPGLGEPPTPAVGPPTMASGGGAPMADNNEGADGVGNAGNMHCSTVLGDRV 240
Qy	241	TTSTRITWALPTYNHLYKQISSASTGASNDNHFGYSTPMGYEDFNHFGHFSRDMQRL 300
Db	241	TTSTRITWALPTYNHLYKQISSASTGASNDNHFGYSTPMGYEDFNHFGHFSRDMQRL 300
Qy	301	INNMGPRKRLNFKLFIQVKEVTTNDGVTTLANNLTITVQVPSDSEYQLPYLSAHQ 360

```

Db      301  INNMGFRPRKLNFKLFNIQVKEVTTNDGVTIANNLSTVQVFSDEYQLPYVLGSAHQ 360
Qy      361  GCLPPPADVFMTPOYGYLLTNNNGSAVGRSSFYCLEYPPSOMLRGNNFTFSYTEEVP 420
Db      361  GCLPPPADVFMTPOYGYLLTNNNGSAVGRSSFYCLEYPPSOMLRGNNFTFSYTEEVP 420
Qy      421  FHSSYAHOSGLDLAMPPLIDQYLYLNRTONOGSAQNKDILFRSGSPAGMSVQPKNMLP 480
Db      421  FHSSYAHOSGLDLAMPPLIDQYLYLNRTONOGSAQNKDILFRSGSPAGMSVQPKNMLP 480
Qy      481  GPCYRQORVSKTIDNNNSNFTWTGASKYMLNGRESIINPGTAMASHKDEDEKFFPMGCV 540
Db      481  GPCYRQORVSKTIDNNNSNFTWTGASKYMLNGRESIINPGTAMASHKDEDEKFFPMGCV 540
Qy      541  MIFGKSAGASNTALDNVMTTDEBEIKATNPVATERFGTAVANFQSSSTDPAIGDVAHMG 600
Db      541  MIFGKSAGASNTALDNVMTTDEBEIKATNPVATERFGTAVANFQSSSTDPAIGDVAHMG 600
Qy      601  ALPGMWQDRDYLQGPIMAKIPHTDGHFHPSPLMGFGKPNPPOLLINRTVPPANPA 660
Db      601  ALPGMWQDRDYLQGPIMAKIPHTDGHFHPSPLMGFGKPNPPOLLINRTVPPANPA 660
Qy      661  EFSATKFAPIITQYSTGQVSEIEMELQKENSKRMPNEVOYTSNYAKSANVDFTVDNNG 720
Db      661  EFSATKFAPIITQYSTGQVSEIEMELQKENSKRMPNEVOYTSNYAKSANVDFTVDNNG 720
Qy      721  YTEPRPIGTRYLTRPL 736
Db      721  YTEPRPIGTRYLTRPL 736

RESULT 2
Q6UC08_9VIRU      PRELIMINARY;  PRT;  736 AA.
AC  06UC08;
DT  05-JUL-2004, integrated into UniProtKB/TrEMBL.
DT  05-JUL-2004, sequence version 1.
DT  07-FEB-2006, entry version 7.
DE  Capsid protein VP1.
GN  Name=cap;
OS  Adeno-associated virus.
OC  Viruses; ssDNA viruses; Parvoviridae; Parvovirinae; Dependovirus;
OC  unclassified Dependovirus.
OX  NCBI_TaxID=272636;
RN  [1]
RP  NUCLEOTIDE SEQUENCE.
RX  PubMed=15163731; DOI=10.1128/JVI.78.12.6381-6388.2004;
RA  Gao G., Vandenberghe L.H., Alvira M.R., Lu Y., Calcedo R., Zhou X.,
RA  Wilson J.M.;
RT  "Clades of Adeno-associated viruses are widely disseminated in human
RT  tissues.";
RL  J. Virol. 78:6381-6388(2004).
CC  Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC  Distributed under the Creative Commons Attribution-NoDerivs license
CC  -----
EMBL; AY530611; AAS99296.1; -; Genomic_DNA.
DR  SMR; Q6UC08; 217-736.
DR  GO; GO:0019028; C:Viral capsid; IEA.
DR  GO; GO:0005198; F:Structural molecule activity; IEA.
DR  InterPro; IPR001403; Parvo_coat.
DR  Pfam; PF00740; Parvo_coat.1.
SQ  SEQUENCE 736 AA; 81317 MW; 373C7C40975CD9BD CRC64;

Query Match      99.5%; Score 3968; DB 2; Length 736;
Best local Similarity 99.3%; Pred. No. 9,4e-250;
Matches 731; Conservative 3; Mismatches 2; Indels 0; Gaps 0;
Qy      1  MAADGYLPDWLEDNLSGIREWMDLKPAPKPKANQKQKODGRGLVLPGYKYLGPENGLD 60
Db      1  MAADGYLPDWLEDNLSGIREWMDLKPAPKPKANQKQKODGRGLVLPGYKYLGPENGLD 60

```

```

Qy      61  KGEPVNAADAAALEHDKAYDQOLKAGDNPLYRNYHADAEFOERLOEDTSFGNIGRAVQ 120
Db      61  KGEPVNAADAAALEHDKAYDQOLKAGDNPLYRNYHADAEFOERLOEDTSFGNIGRAVQ 120
Qy      121  AKKRVLLEPLGLVEGAKTAPGKKRPVEQSPQEBDSSSGIGTKTQOQPAKKRLNFQGTGSE 180
Db      121  AKKRVLLEPLGLVEGAKTAPGKKRPVEQSPQEBDSSSGIGTKTQOQPAKKRLNFQGTGSE 180
Qy      181  SVDDPQPLGEPPTPAAVGPTTMASSGGA PMADNNEGADGVAGASGNWGHDSITWLGKRV 240
Db      181  SVDDPQPLGEPPTPAAVGPTTMASSGGA PMADNNEGADGVAGASGNWGHDSITWLGKRV 240
Qy      241  TTSRTMALPTYNNHLKYKQISSASTGASNDNHFYGSTPMGYDPFNRFHCHFSPPRQRL 300
Db      241  TTSRTMALPTYNNHLKYKQISSASTGASNDNHFYGSTPMGYDPFNRFHCHFSPPRQRL 300
Qy      301  INNMGFRPRKLNFKLFNIQVKEVTTNDGVTIANNLSTVQVFSDEYQLPYVLGSAHQ 360
Db      301  INNMGFRPRKLNFKLFNIQVKEVTTNDGVTIANNLSTVQVFSDEYQLPYVLGSAHQ 360
Qy      361  GCLPPPADVFMTPOYGYLLTNNNGSAVGRSSFYCLEYPPSOMLRGNNFTFSYTEEVP 420
Db      361  GCLPPPADVFMTPOYGYLLTNNNGSAVGRSSFYCLEYPPSOMLRGNNFTFSYTEEVP 420
Qy      421  FHSSYAHOSGLDLAMPPLIDQYLYLNRTONOGSAQNKDILFRSGSPAGMSVQPKNMLP 480
Db      421  FHSSYAHOSGLDLAMPPLIDQYLYLNRTONOGSAQNKDILFRSGSPAGMSVQPKNMLP 480
Qy      481  GPCYRQORVSKTIDNNNSNFTWTGASKYMLNGRESIINPGTAMASHKDEDEKFFPMGCV 540
Db      481  GPCYRQORVSKTIDNNNSNFTWTGASKYMLNGRESIINPGTAMASHKDEDEKFFPMGCV 540
Qy      541  MIFGKSAGASNTALDNVMTTDEBEIKATNPVATERFGTAVANFQSSSTDPAIGDVAHMG 600
Db      541  MIFGKSAGASNTALDNVMTTDEBEIKATNPVATERFGTAVANFQSSSTDPAIGDVAHMG 600
Qy      601  ALPGMWQDRDYLQGPIMAKIPHTDGHFHPSPLMGFGKPNPPOLLINRTVPPANPA 660
Db      601  ALPGMWQDRDYLQGPIMAKIPHTDGHFHPSPLMGFGKPNPPOLLINRTVPPANPA 660
Qy      661  EFSATKFAPIITQYSTGQVSEIEMELQKENSKRMPNEVOYTSNYAKSANVDFTVDNNG 720
Db      661  EFSATKFAPIITQYSTGQVSEIEMELQKENSKRMPNEVOYTSNYAKSANVDFTVDNNG 720
Qy      721  YTEPRPIGTRYLTRPL 736
Db      721  YTEPRPIGTRYLTRPL 736

RESULT 3
O56137_9VIRU      PRELIMINARY;  PRT;  736 AA.
AC  O56137;
DT  01-JUN-1998, integrated into UniProtKB/TrEMBL.
DT  01-JUN-1998, sequence version 1.
DT  07-FEB-2006, entry version 19.
DE  Capsid protein VP1.
OS  Adeno-associated virus 6.
OC  Viruses; ssDNA viruses; Parvoviridae; Parvovirinae; Dependovirus.
OX  NCBI_TaxID=68558;
RN  [1]
RP  NUCLEOTIDE SEQUENCE.
RX  MEDLINE=98080418; PubMed=9420229;
RA  Rutledge B.A., Halbert C.L., Ruesell D.W.;
RT  "Infectious clones and vectors derived from adeno-associated virus
RT  (AAV) serotypes other than AAV type 2.";
RL  J. Virol. 72:309-319(1998).
CC  Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC  Distributed under the Creative Commons Attribution-NoDerivs license
CC  -----
EMBL; AF028704; AAB95450.1; -; Genomic_DNA.
DR  SMR; O56137; 217-736.

```

DR GO: GO:0019028; C:Viral capsid; IEA.
 DR GO: GO:0005198; F:Structural molecule activity; IEA.
 DR InterPro: IPR001403; Parvo_coat.
 DR Pfam: PF00740; Parvo_coat; 1.
 SQ SEQUENCE 736 AA; 81411 MW; 311217A089C565F5 CRC64;

Query Match 99.3%; Score 3963; DB 2; Length 736;
 Best Local Similarity 99.2%; Pred. No. 2e-249;
 Matches 730; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 1 MAADGYPLDMLIEDLSEIGREWMDLKGPAPKPKANQOKODDGRGLVPGYKYLGFNFGLD 60
 DB 1 MAADGYPLDMLIEDLSEIGREWMDLKGPAPKPKANQOKODDGRGLVPGYKYLGFNFGLD 60
 QY 61 KGEVNAADAALAEHDKAYDQOLKAGDNPYLRYNHADAERLQEDTSPFGNIGRAVFQ 120
 DB 61 KGEVNAADAALAEHDKAYDQOLKAGDNPYLRYNHADAERLQEDTSPFGNIGRAVFQ 120
 QY 121 AKKRVLEPLGLVEGATTAEGKKRPVEQSPQEPDSSSGIGTKGQOPAKKRLNFGQTGDS 180
 DB 121 AKKRVLEPLGLVEGATTAEGKKRPVEQSPQEPDSSSGIGTKGQOPAKKRLNFGQTGDS 180
 QY 181 SVDPQPLGEPPTPAVGPPTMASGGAPMADNNEGADGVNAGNMHCOSTWLGDRV 240
 DB 181 SVDPQPLGEPPTPAVGPPTMASGGAPMADNNEGADGVNAGNMHCOSTWLGDRV 240
 QY 241 TTSRTWALPTYNHLYKQISSASTGASNDNHYFGYSTPMGYDPFNRFHCHFSFRDQRL 300
 DB 241 TTSRTWALPTYNHLYKQISSASTGASNDNHYFGYSTPMGYDPFNRFHCHFSFRDQRL 300
 QY 301 INNNMGRPRKLNFKLNIQVKEVTNDGVTIANNLSTVQVFSDEYQLPYLGSAAHQ 360
 DB 301 INNNMGRPRKLNFKLNIQVKEVTNDGVTIANNLSTVQVFSDEYQLPYLGSAAHQ 360
 QY 361 GCLPPFPADVPMIPOGYLTLNNGSOAVGRSFFCYCLEYFSPQMLRTGNFTFSYTFEBVP 420
 DB 361 GCLPPFPADVPMIPOGYLTLNNGSOAVGRSFFCYCLEYFSPQMLRTGNFTFSYTFEBVP 420
 QY 421 FHSSVYAHQSGLDRLMNPIDQYLYLNRTONQSGAQNKDLFSRGSFAGMSVQPKWMLP 480
 DB 421 FHSSVYAHQSGLDRLMNPIDQYLYLNRTONQSGAQNKDLFSRGSFAGMSVQPKWMLP 480
 QY 481 GPCYRQORVSKTKTDNNNSNFTWTGASKYNLNGRESIINPGTAMASHKDEDEKFFPMSCV 540
 DB 481 GPCYRQORVSKTKTDNNNSNFTWTGASKYNLNGRESIINPGTAMASHKDEDEKFFPMSCV 540
 QY 541 MIFGKSAGASNTALDNVMTDEEIKATNPVATERFGTVAVNFQSSSTDPATGDVHAMG 600
 DB 541 MIFGKSAGASNTALDNVMTDEEIKATNPVATERFGTVAVNFQSSSTDPATGDVHAMG 600
 QY 601 ALPGMWQODRDVYLQGPIMAKIPHTDGHFHPSPLMGFGGLKNPPOILLIKNTVPANPPA 660
 DB 601 ALPGMWQODRDVYLQGPIMAKIPHTDGHFHPSPLMGFGGLKNPPOILLIKNTVPANPPA 660
 QY 661 EFSATKFASTTTOYSTGQSVSEIEMELQKENSKRANPEVOYTSNYAKSANVDFTVDNGL 720
 DB 661 EFSATKFASTTTOYSTGQSVSEIEMELQKENSKRANPEVOYTSNYAKSANVDFTVDNGL 720
 QY 721 YTEPRPIGRTYLRPL 736
 DB 721 YTEPRPIGRTYLRPL 736

RESULT 4
 06JUL3_9VIRU PRELIMINARY; PRT: 737 AA.
 AC 06JUL3_9VIRU
 DT 05-JUL-2004, integrated into UniProtKB/TrEMBL.
 DT 07-FEB-2006, entry version 7.
 DE Capsid protein VP1.
 GN Name:cap.
 OS Adeno-associated virus.

OC Viruses; ssDNA viruses; Parvoviridae; Parvovirinae; Dependovirus;
 CC unclassified Dependovirus.
 OX NCBI_TaxID=272636;
 RN [1]
 RP NUCLEOTIDE SEQUENCE.
 RX PubMed=15163731; DOI=10.1128/JVI.78.12.6381-6388.2004;
 RA Gao G., Vandenberghie L.H., Alvira M.R., Lu Y., Calcedo R., Zhou X.,
 RA Wilson J.M.,
 RT "Clades of Adeno-associated viruses are widely disseminated in human
 tissues.";
 RL J. Virol. 78:6381-6388(2004).
 CC -----
 CC Copyrighted by the UniProt Consortium, see <http://www.uniprot.org/terms>
 CC Distributed under the Creative Commons Attribution-NonDerivs License
 CC -----
 DR EMBL: AY530606; AAS99291.1; -; Genomic_DNA.
 DR SMR: 06JUL3; 218-737.
 DR GO: GO:0019028; C:Viral capsid; IEA.
 DR GO: GO:0005198; F:Structural molecule activity; IEA.
 DR InterPro: IPR001403; Parvo_coat.
 DR Pfam: PF00740; Parvo_coat; 1.
 SQ SEQUENCE 737 AA; 81332 MW; 94CED0F3AFDE492 CRC64;

Query Match 97.4%; Score 3883.5; DB 2; Length 737;
 Best Local Similarity 97.4%; Pred. No. 3e-244;
 Matches 718; Conservative 6; Mismatches 12; Indels 1; Gaps 1;

QY 1 MAADGYPLDMLIEDLSEIGREWMDLKGPAPKPKANQOKODDGRGLVPGYKYLGFNFGLD 60
 DB 1 MAADGYPLDMLIEDLSEIGREWMDLKGPAPKPKANQOKODDGRGLVPGYKYLGFNFGLD 60
 QY 61 KGEVNAADAALAEHDKAYDQOLKAGDNPYLRYNHADAERLQEDTSPFGNIGRAVFQ 120
 DB 61 KGEVNAADAALAEHDKAYDQOLKAGDNPYLRYNHADAERLQEDTSPFGNIGRAVFQ 120
 QY 121 AKKRVLEPLGLVEGATTAEGKKRPVEQSPQEPDSSSGIGTKGQOPAKKRLNFGQTGDS 179
 DB 121 AKKRVLEPLGLVEGATTAEGKKRPVEQSPQEPDSSSGIGTKGQOPAKKRLNFGQTGDS 180
 QY 181 SVDPQPLGEPPTPAVGPPTMASGGAPMADNNEGADGVNAGNMHCOSTWLGDRV 239
 DB 181 SVDPQPLGEPPTPAVGPPTMASGGAPMADNNEGADGVNAGNMHCOSTWLGDRV 240
 QY 241 TTSRTWALPTYNHLYKQISSASTGASNDNHYFGYSTPMGYDPFNRFHCHFSFRDQRL 300
 DB 241 TTSRTWALPTYNHLYKQISSASTGASNDNHYFGYSTPMGYDPFNRFHCHFSFRDQRL 300
 QY 301 INNNMGRPRKLNFKLNIQVKEVTNDGVTIANNLSTVQVFSDEYQLPYLGSAAHQ 359
 DB 301 INNNMGRPRKLNFKLNIQVKEVTNDGVTIANNLSTVQVFSDEYQLPYLGSAAHQ 360
 QY 361 GCLPPFPADVPMIPOGYLTLNNGSOAVGRSFFCYCLEYFSPQMLRTGNFTFSYTFEBVP 419
 DB 361 GCLPPFPADVPMIPOGYLTLNNGSOAVGRSFFCYCLEYFSPQMLRTGNFTFSYTFEBVP 420
 QY 421 FHSSVYAHQSGLDRLMNPIDQYLYLNRTONQSGAQNKDLFSRGSFAGMSVQPKWML 479
 DB 421 FHSSVYAHQSGLDRLMNPIDQYLYLNRTONQSGAQNKDLFSRGSFAGMSVQPKWML 480
 QY 481 GPCYRQORVSKTKTDNNNSNFTWTGASKYNLNGRESIINPGTAMASHKDEDEKFFPMSCV 539
 DB 481 GPCYRQORVSKTKTDNNNSNFTWTGASKYNLNGRESIINPGTAMASHKDEDEKFFPMSCV 540
 QY 540 VMI FGKSAGASNTALDNVMTDEEIKATNPVATERFGTVAVNFQSSSTDPATGDVHAM 599
 DB 541 VMI FGKSAGASNTALDNVMTDEEIKATNPVATERFGTVAVNFQSSSTDPATGDVHAM 600
 QY 600 GALPGMWQODRDVYLQGPIMAKIPHTDGHFHPSPLMGFGGLKNPPOILLIKNTVPANPP 659
 DB 601 GALPGMWQODRDVYLQGPIMAKIPHTDGHFHPSPLMGFGGLKNPPOILLIKNTVPANPP 660
 QY 660 AFSATKFASTTTOYSTGQSVSEIEMELQKENSKRANPEVOYTSNYAKSANVDFTVDNGL 719
 DB 660 AFSATKFASTTTOYSTGQSVSEIEMELQKENSKRANPEVOYTSNYAKSANVDFTVDNGL 719

```

Db      661 AERSATKFASTITQYSTGVSVIEIEMELQKENSKRNMPEVQYTSNYAKSASVDFTVDNNG 720
Qy      720 LYTEPRPIGTRYLTRPL 736
      |||||
Db      721 LYTEPRPIGTRYLTRPL 737

RESULT 5
Q6JC12_9VIRU
ID Q6JC12_9VIRU PRELIMINARY; PRT; 736 AA.
AC Q6JC12;
DT 05-JUL-2004, integrated into UniProtKB/TrEMBL.
DT 07-FEB-2006, sequence version 1.
DE Capsid protein VPI.
GN Name=cap;
OS Adeno-associated virus.
OC Viruses; ssDNA viruses; Parvoviridae; Parvovirinae; Dependovirus;
OC unclassified Dependovirus.
OX NCBI_TaxID=272636;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RX PubMed=15163731; DOI=10.1128/JVI.78.12.6381-6388.2004;
RA Gao G., Vandenberghe L.H., Alvira M.R., Lu Y., Calcedo R., Zhou X.,
RT "Clades of Adeno-associated viruses are widely disseminated in human
RT tissues.";
RL J. Virol. 78:6381-6388(2004).
CC -----
CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NonDerivs license
CC -----
DR EMBL; AY530607; AAS9292.1; -; Genomic_DNA.
DR SMR; Q6JC12; 217-736.
DR GO; GO:0019028; C:Viral capsid; IEA.
DR GO; GO:0005198; F:structural molecule activity; IEA.
DR InterPro; IPR001403; Parvo_coat.
DR Pfam; PF00740; Parvo_coat; 1.
SQ SEQUENCE 736 AA; 81364 MW; F2D75A06662E6376 CRC64;

Query Match 97.1%; Score 3873; DB 2; Length 736;
Best Local Similarity 96.9%; Pred. No. 1,5e-243;
Matches 713; Conservative 10; Mismatches 13; Indels 0; Gaps 0;

Qy      1 MAADGYLPMLEEDNLSEGRBEMWDLKPGAKPKANQCKODDGRGLVLPGRKYLGPFNGLD 60
      |||||
Db      1 MAADGYLPMLEEDNLSEGRBEMWDLKPGAKPKANQCKODDGRGLVLPGRKYLGPFNGLD 60
      |||||
Qy      61 KGEPVNAADAALAEHDKAYDOOLKAGDNPYLRVNHADAERFOERLQEDTSFGNLRGAVFQ 120
      |||||
Db      61 KGEPVNAADAALAEHDKAYDOOLKAGDNPYLRVNHADAERFOERLQEDTSFGNLRGAVFQ 120
      |||||
Qy      121 AKKRVLEPLGLVEGAKTAPGKKRPVQSPQEPDSSSGIGKTGOQPAKKRLNFGQTGDS 180
      |||||
Db      121 AKKRVLEPLGLVEGAKTAPGKKRPVQSPQEPDSSSGIGKTGOQPAKKRLNFGQTGDS 180
      |||||
Qy      121 AKKRVLEPLGLVEGAKTAPGKKRPVQSPQEPDSSSGIGKTGOQPAKKRLNFGQTGDS 180
      |||||
Db      121 AKKRVLEPLGLVEGAKTAPGKKRPVQSPQEPDSSSGIGKTGOQPAKKRLNFGQTGDS 180
      |||||
Qy      181 SVDPDQGLGEPAPTPAAVGFPTWASGGGAPMADNNEGADVGNASGNWCHDSTWLGDRVI 240
      |||||
Db      181 SVDPDQGLGEPAPTPAAVGFPTWASGGGAPMADNNEGADVGNASGNWCHDSTWLGDRVI 240
      |||||
Qy      241 TTSTRTALPTNNHLTKQISSASTGASNDNHFGYSTPMGYFPENFCHGFSRDMQRL 300
      |||||
Db      241 TTSTRTALPTNNHLTKQISSASTGASNDNHFGYSTPMGYFPENFCHGFSRDMQRL 300
      |||||
Qy      301 INNNMGRPKRLNFKLNIQVEKVTINDGVTTINNLNLTSTVQVSDSEYQLPYLGSAAQ 360
      |||||
Db      301 INNNMGRPKRLNFKLNIQVEKVTINDGVTTINNLNLTSTVQVSDSEYQLPYLGSAAQ 360
      |||||
Qy      361 GCLPPFPADVFMIPQYGYLTLLNGSQA VGRSSPYCLEYFPBQMLRTGNNFTFSYTFEEVP 420
      |||||
Db      361 GCLPPFPADVFMIPQYGYLTLLNGSQA VGRSSPYCLEYFPBQMLRTGNNFTFSYTFEEVP 420
      |||||
Qy      421 FHSSTASHQSGLDRMLNPLIDQYLYYLNRTQNGSQAQNKDLLFSRGSBPAGNSVQPKMWLP 480
      |||||

```

```

Db      421 FHSSTASHQSGLDRMLNPLIDQYLYYLNRTQNGSQAQNKDLLFSRGSBPAGNSVQPKMWLP 480
      |||||
Qy      481 GPCYRQORVSKTIDNNNSNFTWTGASKYNLNGRESI INPGTAMASHKDEDEKFPWGSV 540
      |||||
Db      481 GPCYRQORVSKTIDNNNSNFTWTGASKYNLNGRESI INPGTAMASHKDEDEKFPWGSV 540
      |||||
Qy      541 MTFKESAGASNTALNNWITDEEIKATNPVATERFGTVANFOSSSTDPAQDVHAMG 600
      |||||
Db      541 MTFKESAGASNTALNNWITDEEIKATNPVATERFGTVANFOSSSTDPAQDVHAMG 600
      |||||
Qy      601 ALPGWMOBRDYLQGPINAKIPHTDGHFPPSPIMGFGHKNPPOILIKNTFVPANPPA 660
      |||||
Db      601 ALPGWMOBRDYLQGPINAKIPHTDGHFPPSPIMGFGHKNPPOILIKNTFVPANPPA 660
      |||||
Qy      661 EFSATKFASTITQYSTGVSVIEIEMELQKENSKRNMPEVQYTSNYAKSASVDFTVDNNG 720
      |||||
Db      661 EFSATKFASTITQYSTGVSVIEIEMELQKENSKRNMPEVQYTSNYAKSASVDFTVDNNG 720
      |||||
Qy      721 LYTEPRPIGTRYLTRPL 736
      |||||
Db      721 LYTEPRPIGTRYLTRPL 736
      |||||

RESULT 6
Q6JC10_9VIRU
ID Q6JC10_9VIRU PRELIMINARY; PRT; 736 AA.
AC Q6JC10;
DT 05-JUL-2004, integrated into UniProtKB/TrEMBL.
DT 07-FEB-2006, sequence version 1.
DE Capsid protein VPI.
GN Name=cap;
OS Adeno-associated virus.
OC Viruses; ssDNA viruses; Parvoviridae; Parvovirinae; Dependovirus;
OC unclassified Dependovirus.
OX NCBI_TaxID=272636;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RX PubMed=15163731; DOI=10.1128/JVI.78.12.6381-6388.2004;
RA Gao G., Vandenberghe L.H., Alvira M.R., Lu Y., Calcedo R., Zhou X.,
RT "Clades of Adeno-associated viruses are widely disseminated in human
RT tissues.";
RL J. Virol. 78:6381-6388(2004).
CC -----
CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NonDerivs license
CC -----
DR EMBL; AY530609; AAS9294.1; -; Genomic_DNA.
DR SMR; Q6JC10; 217-736.
DR GO; GO:0019028; C:Viral capsid; IEA.
DR GO; GO:0005198; F:structural molecule activity; IEA.
DR InterPro; IPR001403; Parvo_coat.
DR Pfam; PF00740; Parvo_coat; 1.
SQ SEQUENCE 736 AA; 81388 MW; BA9463B970028DF0 CRC64;

Query Match 96.6%; Score 3853; DB 2; Length 736;
Best Local Similarity 96.6%; Pred. No. 2,9e-242;
Matches 711; Conservative 9; Mismatches 16; Indels 0; Gaps 0;

Qy      1 MAADGYLPMLEEDNLSEGRBEMWDLKPGAKPKANQCKODDGRGLVLPGRKYLGPFNGLD 60
      |||||
Db      1 MAADGYLPMLEEDNLSEGRBEMWDLKPGAKPKANQCKODDGRGLVLPGRKYLGPFNGLD 60
      |||||
Qy      61 KGEPVNAADAALAEHDKAYDOOLKAGDNPYLRVNHADAERFOERLQEDTSFGNLRGAVFQ 120
      |||||
Db      61 KGEPVNAADAALAEHDKAYDOOLKAGDNPYLRVNHADAERFOERLQEDTSFGNLRGAVFQ 120
      |||||
Qy      121 AKKRVLEPLGLVEGAKTAPGKKRPVQSPQEPDSSSGIGKTGOQPAKKRLNFGQTGDS 180
      |||||
Db      121 AKKRVLEPLGLVEGAKTAPGKKRPVQSPQEPDSSSGIGKTGOQPAKKRLNFGQTGDS 180
      |||||

```

```

QY 181 SVPPQPLGEPAPPAVPTTMAAGGAPMADNNEGADGVGNASGNHCHSTWLGDRYI 240
Db 181 SVPPQPLGEPAPPAVPTTMAAGGAPMADNNEGADGVGNASGNHCHSTWLGDRYI 240
QY 241 TTSRTMALPTYNNHLYKQISSASTGASNDNHYFGYSTPMGCFDNRPHCHSPRDMQRL 300
Db 241 TTSRTMALPTYNNHLYKQISSASTGASNDNHYFGYSTPMGCFDNRPHCHSPRDMQRL 300
QY 301 INNMWGFPRKRLNFKLFNIQVKEVTTNDGVTTIANNLSTVQVPSDSEYQLPYVLGSAHQ 360
Db 301 INNMWGFPRKRLNFKLFNIQVKEVTTNDGVTTIANNLSTVQVPSDSEYQLPYVLGSAHQ 360
QY 361 GCLPPFPADVPMIPOYGYLTINNGSQAVGRSSFYCLEYFPQSOMLRTGNNFTFSYTFEEVP 420
Db 361 GCLPPFPADVPMIPOYGYLTINNGSQAVGRSSFYCLEYFPQSOMLRTGNNFTFSYTFEEVP 420
QY 421 FHSSYAHOSQIDRLMNPILIDQYLYLNRTQNSQAKNDLLFSRGSFAGMSVOPKMWLP 480
Db 421 LHSSCAHSQSDRLMNPILIDQYLYLNRTQNSQAKNDLLFSRGSFAGMSVOPKMWLP 480
QY 481 GPCYRQORVSKTKTDNNNSNFTWTGASKYLNNGRESIINPGTAMASHKDEDEKFFPMGCV 540
Db 481 GPCYRQORVSKTKTDNNNSNFTWTGASKYLNNGRESIINPGTAMASHKDEDEKFFPMGCV 540
QY 541 MIFKESAGASNTALDNVMTDEBEIKATNPVATERFCTVAVNFQSSSTDPAIGVHAMG 600
Db 541 MIFKESAGASNTALDNVMTDEBEIKATNPVATERFCTVAVNFQSSSTDPAIGVHAMG 600
QY 601 ALPGMWQODRDVYLQGPIMAKIPHTDGHFSPMLMGSGGLKNPPQILLIKTVPANPAP 660
Db 601 ALPGMWQODRDVYLQGPIMAKIPHTDGHFSPMLMGSGGLKNPPQILLIKTVPANPAP 660
QY 661 EFSATKPAASFTIQTGVSQVSEIEMELOKENSKRNPEVQYTSNYSASANDFTVDNNG 720
Db 661 EFSATKPAASFTIQTGVSQVSEIEMELOKENSKRNPEVQYTSNYSASANDFTVDNNG 720
QY 721 YTEPRPIGTRYLTRPL 736
Db 721 YTEPRPIGTRYLTRPL 736

RESULT 7
056139_9VIRU PRELIMINARY; PRT; 736 AA.
AC 056139;
DT 01-JUN-1998, integrated into UniProtKB/TrEMBL.
DT 01-JUN-1998, sequence version 1.
DT 07-FEB-2006, entry version 20.
DE Capsid protein VP1.
OS Adeno-associated virus 3B.
OC Viruses; ssDNA viruses; Parvoviridae; Parvovirinae; Dependovirus.
CX NCBI_TaxID=68742;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=98080418; PubMed=9420229;
RA Rutledge E.A., Halbert C.L., Russell D.W.;
RT "Infectious clones and vectors derived from adeno-associated virus
(J. Virol. 72:309-319 (1998))."
CC -----
CC Distributed under the UniProt Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NonDerivs license
CC -----
DR EMBL; AF028705; AAB95452.1; -; Genomic_DNA.
DR SMR; 056139; 217-736.
DR GO; GO:0019028; C:viral capsid; IEA.
DR GO; GO:0005158; F:structural molecule activity; IEA.
DR InterPro; IPR001403; Parvo_coat.
DR Pfam; PF00740; Parvo_coat; 1.
SQ SEQUENCE 736 AA; 81907 MW; D52331AD5F0D70F CRC64;

Query March 86.0%; Score 3511; DB 2; Length 736;
Best Local Similarity 86.8%; Pred. NO. 5.7e-220;

```

```

Matches 640; Conservative 38; Mismatches 57; Indels 2; Gaps 2;
QY 1 MAAGYLPDWLEDNLSSEIGIREMMDLKPGAPRKANQKQODGRLVLPYKYLGPNGLD 60
Db 1 MAAGYLPDWLEDNLSSEIGIREMMDLKPGAPRKANQKQODGRLVLPYKYLGPNGLD 60
QY 61 KGEPMNADAAALAEHDKAYDQOLKAGDNPYLRYNHADAEPQERLEOEDTSPFGNLCRAVFO 120
Db 61 KGEPMNADAAALAEHDKAYDQOLKAGDNPYLRYNHADAEPQERLEOEDTSPFGNLCRAVFO 120
QY 121 AKKRLPEPLGVEBAGAKTAPGKKRPVEQSPQESSGIGTKGOQPAKRLNFGQTGSE 180
Db 121 AKKRLPEPLGVEBAGAKTAPGKKRPVDSPOSSSGVGSQKQAPKRLNFGQTGSE 180
QY 181 SVPPQPLGEPAPPAVPTTMAAGGAPMADNNEGADGVGNASGNHCHSTWLGDRYI 240
Db 181 SVPPQPLGEPAPPAVPTTMAAGGAPMADNNEGADGVGNASGNHCHSTWLGDRYI 240
QY 241 TTSRTMALPTYNNHLYKQISSASTGASNDNHYFGYSTPMGCFDNRPHCHSPRDMQRL 300
Db 241 TTSRTMALPTYNNHLYKQISSASTGASNDNHYFGYSTPMGCFDNRPHCHSPRDMQRL 300
QY 301 INNMWGFPRKRLNFKLFNIQVKEVTTNDGVTTIANNLSTVQVPSDSEYQLPYVLGSAHQ 360
Db 301 INNMWGFPRKRLNFKLFNIQVKEVTTNDGVTTIANNLSTVQVPSDSEYQLPYVLGSAHQ 360
QY 361 GCLPPFPADVPMIPOYGYLTINNGSQAVGRSSFYCLEYFPQSOMLRTGNNFTFSYTFEEVP 420
Db 361 GCLPPFPADVPMIPOYGYLTINNGSQAVGRSSFYCLEYFPQSOMLRTGNNFTFSYTFEEVP 420
QY 421 FHSSYAHOSQIDRLMNPILIDQYLYLNRTQNSQAKNDLLFSRGSFAGMSVOPKMWLP 479
Db 421 FHSSYAHOSQIDRLMNPILIDQYLYLNRTQNSQAKNDLLFSRGSFAGMSVOPKMWLP 479
QY 480 GPCYRQORVSKTKTDNNNSNFTWTGASKYLNNGRESIINPGTAMASHKDEDEKFFPMGCV 539
Db 480 GPCYRQORVSKTKTDNNNSNFTWTGASKYLNNGRESIINPGTAMASHKDEDEKFFPMGCV 539
QY 540 VMIKESAGASNTALDNVMTDEBEIKATNPVATERFCTVAVNFQSSSTDPAIGVHAM 599
Db 540 VMIKESAGASNTALDNVMTDEBEIKATNPVATERFCTVAVNFQSSSTDPAIGVHAM 599
QY 600 GALPGMWQODRDVYLQGPIMAKIPHTDGHFSPMLMGSGGLKNPPQILLIKTVPANPAP 659
Db 600 GALPGMWQODRDVYLQGPIMAKIPHTDGHFSPMLMGSGGLKNPPQILLIKTVPANPAP 659
QY 660 AFSATKPAASFTIQTGVSQVSEIEMELOKENSKRNPEVQYTSNYSASANDFTVDNNG 719
Db 660 AFSATKPAASFTIQTGVSQVSEIEMELOKENSKRNPEVQYTSNYSASANDFTVDNNG 719
QY 720 LYTEPRPIGTRYLTRPL 736
Db 720 LYTEPRPIGTRYLTRPL 736

RESULT 8
056311_9VIRU PRELIMINARY; PRT; 736 AA.
AC 056311;
DT 01-NOV-1996, integrated into UniProtKB/TrEMBL.
DT 01-NOV-1996, sequence version 1.
DT 07-FEB-2006, entry version 24.
DE Capsid protein.
OS Adeno-associated virus - 3.
OC Viruses; ssDNA viruses; Parvoviridae; Parvovirinae; Dependovirus.
CX NCBI_TaxID=46350;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RX STRAIN=3H;
RA Muramatsu S., Mizukami H., Young N.S., Brown K.E.;
RT "Nucleotide sequencing and generation of an infectious clone of adeno-
associated virus 3.";

```

Accession	Query Match	Score	DB	Length
CC	Query Match	87.6%	Score 3494;	DB 2; Length 736;
CC	Best Local Similarity	86.4%	Pred. No. 7,3e-219;	
CC	Matches 637; Conservative	39;	Mismatches 59;	Indels 2; Gaps 2;
CC	Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms			
CC	Distributed under the Creative Commons Attribution-NonCommercial License			
CC	-----			
DR	EMBL: U48704; AAC55049.1; -; Genomic_DNA.			
DR	SMR: Q65311; 217-736.			
DR	GO: 0019028; C: viral capsid; IEA.			
DR	GO: 0005198; F: structural molecule activity; IEA.			
DR	InterPro: IPR01403; Parvo_coat.			
DR	Pfam: PF00740; Parvo_coat; 1.			
DR	SEQUENCE 736 AA; B1660 MW; APLERF47B5C67A10 CnC64;			
QY	1 MAADSYLPLDMIEDNISEGIREWMDLKPAAPKPNQOQODDGRGIVLPGYKYLLPFGNLD	60		
Db	1 MAADSYLPLDMIEDNISEGIREWMDLKPAAPKPNQOQODDGRGIVLPGYKYLLPFGNLD	60		
QY	61 KGEPNADAAALHEDKAYDQOLKAGDPVYLRVYHADAEPOERLOEDTSPGANGRAVQ	120		
Db	61 KGEPNADAAALHEDKAYDQOLKAGDPVYLRVYHADAEPOERLOEDTSPGANGRAVQ	120		
QY	121 AKKRVLEPLGLVEEGAKTAPGKRPEVDSPOEDSSSGIGTGOQPAKKRLNFGQTGSE	180		
Db	121 AKKRVLEPLGLVEEGAKTAPGKRPEVDSPOEDSSSGIGTGOQPAKKRLNFGQTGSE	180		
QY	121 AKKRILPEPLGLVEEAAKTAAPGKGAVDQSPQEDSSSGVGSKGPAKKRLNFGQTGSE	180		
Db	121 AKKRILPEPLGLVEEAAKTAAPGKGAVDQSPQEDSSSGVGSKGPAKKRLNFGQTGSE	180		
QY	181 SVDPDQPLGEPAPATSPAGSTNMASSGGAPMADNNEGADVGNSSGNNHCDSQWLGDRVI	240		
Db	181 SVDPDQPLGEPAPATSPAGSTNMASSGGAPMADNNEGADVGNSSGNNHCDSQWLGDRVI	240		
QY	241 TTSTETMALPTYNMLLYKOISSASTGASNDNHYFGYSTPWGYPFENRPFHCFSPDMQRL	300		
Db	241 TTSTETMALPTYNMLLYKOISSASTGASNDNHYFGYSTPWGYPFENRPFHCFSPDMQRL	300		
QY	241 TTSTETMALPTYNMLLYKOISSASTGASNDNHYFGYSTPWGYPFENRPFHCFSPDMQRL	299		
Db	241 TTSTETMALPTYNMLLYKOISSASTGASNDNHYFGYSTPWGYPFENRPFHCFSPDMQRL	299		
QY	301 INNNNGFPRKRLNFKLFIQVKEVTTNDGVTTIANNLSTVQVFSDESQOLPYVLGSAHQ	360		
Db	301 INNNNGFPRKRLNFKLFIQVKEVTTNDGVTTIANNLSTVQVFSDESQOLPYVLGSAHQ	360		
QY	300 INNNNGFPRKRLNFKLFIQVKEVTTNDGVTTIANNLSTVQVFSDESQOLPYVLGSAHQ	359		
Db	300 INNNNGFPRKRLNFKLFIQVKEVTTNDGVTTIANNLSTVQVFSDESQOLPYVLGSAHQ	359		
QY	361 GCLPPPADVPMIPQGYTLTLNNGSQAVGRSFFCLETFPFGOMLRTGNNFTPTSTFEVPE	420		
Db	361 GCLPPPADVPMIPQGYTLTLNNGSQAVGRSFFCLETFPFGOMLRTGNNFTPTSTFEVPE	420		
QY	360 GCLPPPADVPMIPQGYTLTLNNGSQAVGRSFFCLETFPFGOMLRTGNNFTPTSTFEVPE	419		
Db	360 GCLPPPADVPMIPQGYTLTLNNGSQAVGRSFFCLETFPFGOMLRTGNNFTPTSTFEVPE	419		
QY	421 FHSSTVAHSQSLDRNLNPLIDQYLLYLNRTQ-NQSGAQNKKLLFSRGSPPAGMSVQPKWL	479		
Db	421 FHSSTVAHSQSLDRNLNPLIDQYLLYLNRTQ-NQSGAQNKKLLFSRGSPPAGMSVQPKWL	479		
QY	420 FHSSTVAHSQSLDRNLNPLIDQYLLYLNRTQ-NQSGAQNKKLLFSRGSPPAGMSVQPKWL	479		
Db	420 FHSSTVAHSQSLDRNLNPLIDQYLLYLNRTQ-NQSGAQNKKLLFSRGSPPAGMSVQPKWL	479		
QY	480 PGPCCYRQORVSKTKTDNNNSNFTWTGASKYMLNGRESILNPGTAVASHKDDDEKFFPMSSG	539		
Db	480 PGPCCYRQORVSKTKTDNNNSNFTWTGASKYMLNGRESILNPGTAVASHKDDDEKFFPMSSG	539		
QY	480 PGPCCYRQORVSKTKTDNNNSNFTWTGASKYMLNGRESILNPGTAVASHKDDDEKFFPMSSG	539		
Db	480 PGPCCYRQORVSKTKTDNNNSNFTWTGASKYMLNGRESILNPGTAVASHKDDDEKFFPMSSG	539		
QY	540 VMIPEGKASGASNTLADNMVITDEBEILATNPVATERGTYAVANFQSSSTDPATGDVHAM	599		
Db	540 VMIPEGKASGASNTLADNMVITDEBEILATNPVATERGTYAVANFQSSSTDPATGDVHAM	599		
QY	540 NLIPEKEGTTASNAELDMVMITDEBEIRTTNVPVATEQYGTAVANNIQQSSNTAPTTGTVAHQ	599		
Db	540 NLIPEKEGTTASNAELDMVMITDEBEIRTTNVPVATEQYGTAVANNIQQSSNTAPTTGTVAHQ	599		
QY	600 GALPGMWMODRNVLYQSPIMAKIPIHTDGHFHSPLMGAFGLKNPPOILLINTPVPANP	659		
Db	600 GALPGMWMODRNVLYQSPIMAKIPIHTDGHFHSPLMGAFGLKNPPOILLINTPVPANP	659		
QY	600 GALPGMWMODRNVLYQSPIMAKIPIHTDGHFHSPLMGAFGLKNPPOILLINTPVPANP	659		
Db	600 GALPGMWMODRNVLYQSPIMAKIPIHTDGHFHSPLMGAFGLKNPPOILLINTPVPANP	659		
QY	660 AEFASATKASFTQYSTQGVSEIEMELQKENSKKMNEVOYTSNYASAVNDFTVANG	719		
Db	660 AEFASATKASFTQYSTQGVSEIEMELQKENSKKMNEVOYTSNYASAVNDFTVANG	719		
QY	660 TTFSPAKFASFTQYSTQGVSEIEMELQKENSKKMNEVOYTSNYASAVNDFTVANG	719		
Db	660 TTFSPAKFASFTQYSTQGVSEIEMELQKENSKKMNEVOYTSNYASAVNDFTVANG	719		
QY	720 LYTEPRPIGTRVLYRLP 736			
Db	720 LYTEPRPIGTRVLYRLP 736			
QY	720 VYSEPRPIGTRVLYRLP 736			
Db	720 VYSEPRPIG			

AC 0808Y3: 2003, integrated into UniProtKB/TrEMBL.
 DT 01-JUN-2003, sequence version 1.
 DT 07-FEB-2006, entry version 11.
 DE Capsid protein.
 GN Name=VP1;
 OS Non-human primate Adeno-associated virus.
 OC Viruses; ssDNA viruses; Parvoviridae; Parvovirinae; Dependovirus;
 OC unclassified Dependovirus.
 OX NCBI_TaxID=226582;
 RN [1]
 RP NUCLEOTIDE SEQUENCE.
 RX MEDLINE=22631657; PubMed=12716974; DOI=10.1073/pnas.0937739100;
 RA Gao G., Alivira M.R., Somanathan S., Lu Y., Vandenberghe L.H.,
 RA Rux J.J., Calcedo R., Samniguel J., Abbas Z., Wilson J.M.;
 RT "Adeno-associated viruses undergo substantial evolution in primates
 RT during natural infections";
 RL Proc. Natl. Acad. Sci. U.S.A. 100:6081-6086(2003).

 CC Copyrighted by the UniProt Consortium, see <http://www.uniprot.org/terms>
 CC Distributed under the Creative Commons Attribution-NonDerivs License
 CC -----
 DR EMBL: AY242997; AAC8183.1; -; Genomic_DNA.
 DR SMR: 0808Y3: 217-736.
 DR GO: GO:0019028; C:Viral capsid; IEA.
 DR GO: GO:0005198; F:Structural molecule activity; IEA.
 DR InterPro: IPR001403; Parvo_coat.
 DR Pfam: PF00740; Parvo_coat; 1.
 DR SQR SEQUENCE 736 AA; 81370 MW; 87D05047FEFA14C CRC64;

 Query Match 87.1%; Score 3473; DB 2; Length 736;
 Best local similarity 85.6%; Pred. No. 1.7e-217;
 Matches 631; Conservative 44; Mismatches 60; Indels 2; Gaps 2;

 QY 1 MAADGYLPDWMLEDNLSEGIREFMMDLKPQAPKPKANQOKQDDGRGLVPGYKYLGPFNGLD 60
 DB 1 MAADGYLPDWMLEDNLSEGIREFMMDLKPQAPKPKANQOKQDDGRGLVPGYKYLGPFNGLD 60
 QY 61 KGEPPNADAALBHDKAYDQOLKAGDNPYLRYNHADAFOERLOEDTSPGNGIGRAVFO 120
 DB 61 KGEPPNADAALBHDKAYDQOLKAGDNPYLRYNHADAFOERLOEDTSPGNGIGRAVFO 120
 QY 121 AKKRVLEPLGLVERGAKTAPGKKRPVEGSPQEPDSSSGIGRTGGQPAKKRLNFGQTGSE 180
 DB 121 AKKRVLEPLGLVERGAKTAPGKKRPVEGSPQEPDSSSGIGRTGGQPAKKRLNFGQTGSE 180
 QY 121 AKKRVLEPLGLVERGAKTAPGKKRPVEGSPQEPDSSSGIGRTGGQPAKKRLNFGQTGSE 180
 DB 121 AKKRVLEPLGLVERGAKTAPGKKRPVEGSPQEPDSSSGIGRTGGQPAKKRLNFGQTGSE 180
 QY 181 SVPDQPLGEPBPAPAAVGPPTTMSGGGAPMADNNEGADGVGNAGNNHCDSTWLGDVYI 240
 DB 181 SVPDQPLGEPBPAPAAVGPPTTMSGGGAPMADNNEGADGVGNAGNNHCDSTWLGDVYI 240
 QY 241 TTSFTFTMALPTYNHNLKQIISASTGAS-NNNHYFGYSTPWGYPDFNRHCHFSBDMQR 299
 DB 241 TTSFTFTMALPTYNHNLKQIISASTGAS-NNNHYFGYSTPWGYPDFNRHCHFSBDMQR 299
 QY 300 LNNWNGFRPKRLNFKLFNIOYKEVTTNDGVTTIANNLTSVVOYVDSSEYQLPYVLGSAH 359
 DB 300 LNNWNGFRPKRLNFKLFNIOYKEVTTNDGVTTIANNLTSVVOYVDSSEYQLPYVLGSAH 359
 QY 301 LIINNMGFRPKRLNFKLFNIOYKEVTTNEGKTTIANNLTSVVOYFTDSEYQLPYVLGSAH 360
 DB 301 LIINNMGFRPKRLNFKLFNIOYKEVTTNEGKTTIANNLTSVVOYFTDSEYQLPYVLGSAH 360
 QY 360 OGCLPPFPADVPMTPQYCVTLTNNGSQAVGSSPFCLEFPFSOMLRTNNFFPSTFEEV 419
 DB 360 OGCLPPFPADVPMTPQYCVTLTNNGSQAVGSSPFCLEFPFSOMLRTNNFFPSTFEEV 419
 QY 420 PFHSSYASQSILDRIMNPLIDQYLYLLRKTQNGSSAQNKDLFSRGSPPAGMSVOPKWL 479
 DB 420 PFHSSYASQSILDRIMNPLIDQYLYLLRKTQNGSSAQNKDLFSRGSPPAGMSVOPKWL 479
 QY 421 PFHSSYASQSILDRIMNPLIDQYLYLLRKTQNGSSAQNKDLFSRGSPPAGMSVOPKWL 479
 DB 421 PFHSSYASQSILDRIMNPLIDQYLYLLRKTQNGSSAQNKDLFSRGSPPAGMSVOPKWL 479
 QY 480 PGPCYRQGRVSKTIDNNNSNFTWTGASKYNLNGRESILNFGTAMASHKDEDEKFFPMSSG 539
 DB 480 PGPCYRQGRVSKTIDNNNSNFTWTGASKYNLNGRESILNFGTAMASHKDEDEKFFPMSSG 539
 QY 540 VMIFKESAGASNTALDNVMTITDEEIKATNPVATERGTYAVAVNFQSSSTDPATGDVYAM 599
 DB 540 VMIFKESAGASNTALDNVMTITDEEIKATNPVATERGTYAVAVNFQSSSTDPATGDVYAM 599
 QY 540 VLIFFKQAGNDGVDSYOVLLITDEEIKATNPVATERGTYAVAVNFQSSSTDPATGDVYAM 599
 DB 540 VLIFFKQAGNDGVDSYOVLLITDEEIKATNPVATERGTYAVAVNFQSSSTDPATGDVYAM 599

QY	600	GALPEMWAODBDVYLQSGITWAKITPHTDGHFHSPLMGGRGLKNPPOLIKNTPTVPANP	653
DB	600	GVLPPEMVAQNKNRVYLQSGITWAKITPHTDGHFHSPLMGGLGHPPOILLIKNTPTVPANP	659
QY	660	AEFSAATKASFSTQYSTGCVSVSEIEMELQKENS KAMNEVOYTSNYASAVNDFTVDNNG	719
DB	660	LTFNQAKLNSFITQYSTGCVSVSEIEMELQKENS KAMNEIQTSTNYKSTNVDFAVNNEG	719
QY	720	LYTEPRPIGTRLYTRPL 736	
DB	720	YVSEBRPIGTRLYTRNL 736	
RESULT 10			
	067008_9VIRU	PRELIMINARY; PRT; 735 AA.	
ID	067008_9VIRU		
AC	067008;		
DT	11-OCT-2004, integrated into UniProtKB/TrEMBL.		
DT	11-OCT-2004, sequence version 1.		
DT	07-FEB-2006, entry version 8.		
DE	Capid protein VP1.		
GN	Name=capd;		
OS	Adeno-associated virus.		
OC	Virusae; ssDNA viruses; Parvoviridae; Parvovirinae; Dependovirus;		
OC	Unclassified Dependovirus.		
OX	NCBI_TaxID=272635;		
RN	[1]		
RP	NUCLEOTIDE SEQUENCE.		
RC	STRAIN=Hu. S17;		
RA	PubMed16282478; DOI=10.1128/JVI.79.23.14781-14792.2005;		
RT	Chen C.L., Jensen R.L., Schnepf B.C., Connell M.J., Shell R.,		
RA	Sferra T.J., Bartlett J.S., Clark K.R., Johnson P.R.,		
RT	"Molecular characterization of adeno-associated viruses infecting		
RT	children";		
RL	J. Virol. 79:14781-14792 (2005).		
CC	-----		
CC	Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms		
CC	Distributed under the Creative Commons Attribution-NonDerivs License		
CC	-----		
DR	EMBL; AY695376; AAU05370.1; -, Genomic_DNA.		
DR	SMR; Q67008; 217-735.		
DR	GO; GO:0019028; C:intracellular capsid; IEA.		
DR	GO; GO:0005198; F:structural molecule activity; IEA.		
DR	InterPro; IPR001403; Parvoviridae coat.		
DR	Pfam; PF00740; Parvoviridae coat; 1.		
SO	SEQUENCE 735 AA; 81888 MW; 256BFB955F1C9925 CRC64;		
Query Match			
	87.0%; Score 3469.5; DB 2; Length 735;		
	Best Local Similarity 85.1%; Pred. No. 2.9e-217;		
Matches	626; Conservative 47; Mismatches 62; Indels 1; Gaps 1;		
QY	1	MAADGYLEPDMLEDNISSEIGREKMDLKPGAPKPKANQKQKODGRLVLPFYKYLGPFGNGLD	60
DB	1	MAAGGYLPDMLEDTLSEGIQRMWMLKPGPPPKPKPAERHKDPSRGILVLPGYKYLGPFGNGLD	60
QY	61	KGEPPNADAALAEHDKAYDQOLKAGNDNPYLRYNADAEFOERLOEDTSPGANGRAYVFO	120
DB	61	KGEPPNEDAALAEHDKAYDROLDSGNDPPLYLTNTADAEFOERLEDTSPGANGRAYVFO	120
QY	121	AKKRVLEPLGIVEEGAKTAPGKKRPEVDSPOEPDSSSGIGTKGQPAKKRLNFGQTGDS	180
DB	121	AKKRVLEPLGIVEEVPKTAAPGKKRPEVHSPAPEDSSSGTGKSGQPAKKRLNFGQTGDS	180
QY	181	SVPDPQPLGEPPTATPAVGPPTTMASGGGAPMADNNEGADGVGNASGNMHCSTNLGDRVY	240
DB	181	SVPDPQPLGEPPEPAATPSLGTSTMASGGGAPVADNNEGADGVGNSSGNMHCDSQNLGDRVY	240
QY	241	TTSTRTMALPTYNHNLVKQISASTGASGANDNHYFGYSTPMGXFDPNRFCHFSPRDMQRL	300
DB	241	TTSTRTMALPTYNHNLVKQISSQS-GASNDNHYFGYSTPMGXFDPNRRHCHFSPRDMQRL	299
QY	301	INNMGFRPKRLNFKLFENIYQKEVTTNDQVTTIANLSTVQVFSDSSEYQLPYVLGSAHQ	360

Db	300	INNNGFPPKRLNEFLFNIQVEKVTQNDGTTIANNLSTVQVFTDSEYQLPVLGSAHQ	359
Qy	361	GCLPEPADVEMIPQYGYLTLLNNGSQAVGRSSFCLEYEPSSOMLRGTNNFTESYTEEYP	420
Db	360	GCLPEPADVEMVPPQYGYLTLLNNGSQAVGRSSFCLEYEPSSOMLRGTNNFTESYTEEDVP	419
Qy	421	FHSSYAHSQSLDRMLNPLIDQYLYLTARTQNGSQAONKDLLFSRGS PAMSVQPKWLP	480
Db	420	FHSSYAHSQSLDRMLNPLIDQYLYLTARTQNGSAGTQOSRLLFSSQGPSMSIAQAKWLP	479
Qy	481	GPCRHOQVSKTKTDNNNSNFTWTGASKYNLNGRESIIINPGTAMASHKDDKDFPMSGV	540
Db	480	GPCYHOQRLSKQAMENNNNSNFPWTATATKXHLNGRSLVNPGRPMASHKDDKDFPMSGT	539
Qy	541	MIFGESASGNLTLDNMTITDEEIEIKATNPVATERFGVVAANFOSSSTDPATGADVHANG	600
Db	540	LIFGKEGTNNANNAELENMTIDEEIEIKRTNPVATQYGVSNLNLOSNAPSGTGYVNHQG	599
Qy	601	ALPGWVMOODRVYLQGPIMAKIPHNDGHFHPSPPLMGFGGLKNNPPOILLIKNTPVVPANPPA	660
Db	600	ALPGWVMOODRVYLQGPIMAKIPHNDGHFHPSPPLMGFGGLKNNPPOILLIKNTPVVPANPP	659
Qy	661	EPSATKPAASFITQYSTGVQSVIEIWELOKENSKRNPVEQYTSNTAKSANVDFTVUNGL	720
Db	660	NFSAKAPFASFITQYSTGVQSVIEIWELOKENSKRNPVEIQYTSNYKSNVVDFTVDNGV	719
Qy	721	YTERPRIGTRYLTPRL	736
Db	720	YSEPRIGTRYLTPRL	735

```

RESULT 11
ID Q6JBZ7_V91RU PRELIMINARY; PRT; 735 AA.
AC Q6JBZ7;
DT 05-JUL-2004, integrated into UniProtKB/TrEMBL.
DT 05-JUL-2004, sequence version 1.
DT 07-FEB-2006, entry version 8.
DE Capsid protein VP1.
GN Name=caps;
OS Adeno-associated virus.
OC Viruses; ssDNA viruses; Parvoviridae; Dependovirus;
OC unclassified dependovirus.
OX NCBI_TaxID=272636;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RX PubMed:15163731; DOI=10.1128/JVI.78.12.6381-6388.2004;
RA Gao G., Vandenberghe L.H., Alviria M.R., Lu Y., Calcedo R., Zhou X.,
RA Wilson J.M.;
RT "Clades of Adeno-associated viruses are widely disseminated in human
tissues.";
RL J. Virol. 78:6381-6388(2004).
-----
CC CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC CC Distributed under the Creative Commons Attribution-NonCommercial license
CC CC -----
DR EMBL; AY530622; AAS99307.1; -; Genomic_DNA.
DR SMR; Q6JBZ7; 217-735.
DR GO; GO:0019028; C:viral capsid; IEA.
DR GO; GO:0005198; F:structural molecule activity; IEA.
DR InterPro; IPR001403; Parvo_coat.
DR Pfam; PF00740; Parvo_coat; 1.
SQ SEQUENCE 735 AA; 81948 MW; BBA318E904E0BF0A CRC64;
-----
Query Match 86.7%; Score 3456.5; DB 2; Length 735;
Best Local Similarity 84.5%; Pred. No. 2e-216;
Matches 622; Conservative 49; Mismatches 64; Indels 1; Gaps 1;
-----
QY 1 MAADGYLPDMLEDTLSEGIKRWMDLKGAPKPKANOOKODDGRGVLPGYKYLGPFGNLD 60
db 1 MAADGYLPDMLEDTLSEGIKRWMDLKGAPKPKAERKHDSSRLGLVLPGYKYLGPFGNLD 60

```


Query Match	Best Local Similarity	Matches	632;	Conservative	48;	Mismatches	55;	Indels	3;	Gaps	3;
DE	Capid protein VPI.										
ON	Name=cap;										
OS	Adeno-associated virus.										
OC	Virusess; ssDNA viruses; Parvoviridae; Parvovirinae; Dependovirus;										
CC	unclassified Dependovirus.										
OX	NCBI_TaxId=272636;										
RN	[1]										
RP	NCBI_EOTIDE SEQUENCE.										
RX	PubMed=15163731; DOI=10.1128/JVI.78.12.6381-6388.2004;										
RA	Gao G., Vandenberghe L.H., Alviria M.R., Lu Y., Calcedo R., Zhou X.,										
RA	Johnson J.M.;										
RT	"Clades of Adeno-associated viruses are widely disseminated in human										
RT	tissues."										
RL	J. Virol. 78:6381-6388 (2004).										
CC	Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms										
CC	distributed under the Creative Commons Attribution-NonDerivs License										
CC											
DR	EMBL; AY530561; AAS99246.1; -; Genomic_DNA.										
DR	SMR; Q6JC58; 218-737.										
DR	GO; GO:0019028; C:viral capsid; IEA.										
DR	GO; GO:0005198; F:structural molecule activity; IEA.										
DR	InterPro: IPR01403; Parvo_coat.										
DR	Pfam: PF00740; Parvo_coat_1.										
DR	SEQUENCE 737 Aa; 81486 MW; EBF685A99F07CB5 CRC64;										
QY	1	MAADGYLPDWLIEDNISEGIREWMWIKPGAPKPKANQKODDGRGLVLPGYKYLGFNGLD	60								
DB	1	MAADGYLPDWLIEDNISEGIREWMWIKPGAPKPKANQKODDGRGLVLPGYKYLGFNGLD	60								
QY	61	KGEPIPNADDAALHEHDKAYDQOLKAGDNPYLRYNHADAEPQERLOEDPESFGNIGRAVFO	120								
DB	61	KGEPIPNADDAALHEHDKAYDQOLKAGDNPYLRYNHADAEPQERLOEDPESFGNIGRAVFO	120								
QY	121	AKKRVLEPLGLVEEGAKTAPGKKRPVEQSPQ-EPDSSGIGTGQOQPKKLNFGQGTDS	179								
DB	121	AKKRVLEPLGLVEEGAKTAPGKKRPVEQSPQSPDSSTGIGTKGQOQPKKLNFGQGTDS	180								
QY	180	ESVPDPQIGEPAPAGPSGLSGGTMAAGGAPADNNKADGAGNAGSMWCHDSITWLGDRV	239								
DB	181	ESVPDPQIGEPAPAGPSGLSGGTMAAGGAPADNNKADGAGNAGSMWCHDSITWLGDRV	240								
QY	240	ITTSRTTALPTYNHNLKYOISASSTAGSNDNHFGYSTPMGYDPFNFHCHFSRDMQR	299								
DB	241	ITTSRTTALPTYNHNLKYOISASSTAGSNDNHFGYSTPMGYDPFNFHCHFSRDMQR	300								
QY	300	LINNMGGRPKRLNFKLFNIOVKETTNDGVTTLANNLTSTVVOVPSDSEYOLPYVLGSAH	359								
DB	301	LINSWGWGRPKKLNFKLFNIOVKETTNDGVTTLANNLTSTVVOVPSDSEYOLPYVLGSAH	360								
QY	360	QGCLPFPFADVFMIPOYGYLTLNNGSQAVGRSSPFYCLEYFPFSOMLRTNNFTFSTYEEV	419								
DB	361	QGCLPFPFADVFMIPOYGYLTLNNGSQAVGRSSPFYCLEYFPFSOMLRTNNFTFSTYEEV	420								
QY	420	PFHSSYHSSQSLDRIMNPLIDQYLYLNRTO-NOSGAHONKOLLFSRSGPAGMSVOPKRW	478								
DB	421	PFHSSYHSSQSLDRIMNPLIDQYLYLNRTO-NOSGAHONKOLLFSRSGPAGMSVOPKRW	480								
QY	479	LPGCYRQORRKTCTDNNNSNFMPTGASKYNLNREIINPGRAMSHKDDKQFPMS	538								
DB	481	LPGCYRQORRKTCTDNNNSNFMPTGASKYNLNREIINPGRAMSHKDDKQFPMS	540								
QY	539	GVMIFGKESAGASNTALDNWMITDEEELKATNPVATERFGVAIVAFOSSTDPATGCVHA	598								
DB	541	GVLIFGKIGKA-ANKTTLENTLVMTNEEBELRPINPVAITEYGVVSSNLQAANAAQTQVYNN	599								
QY	599	MGALPKNWQDRDVLQGPYIAKIPHTDGHFHPSPFLMGFGKLPKPPDILIKATVPANP	658								
DB	600	QGALPKNWQDRDVLQGPYIAKIPHTDGHFHPSPFLMGFGKLPKPPDILIKATVPANP	659								

QY	659	PAEFSATPEASITTOYSTGCVSELEIEMELQENSRKRNPEUYOYSTNVAKSANVDEVTQNN	718
DB	660	PEVFPAPAFASPTIYOSTGVSELEIEMELQENSRKRNPEIYQITNPFQKQGVDAVDSQ	719
QY	719	GLYTEPRPIGRYLTPL 736	
DB	720	GVYSEPRPIGRYLTPL 737	
		RESULT 14	
		06JC28 9VIRU PRELIMINARY; PRT: 735 AA.	
AC	06JC28_		
DT	05-JUL-2004,	integrated into UniProtKB/TrEMBL.	
DT	07-FEB-2006,	entry version 1.	
DE	Capid protein Vp1.		
GN	Name:capid		
OS	Adeno-associated virus.		
OC	Viruses; ssDNA viruses; Parvoviridae; Parvovirinae; Dependovirus;		
OX	NCBI_TaxID=272636;		
RN	[1]		
RP	NUCLEOTIDE SEQUENCE.		
RX	PubMed15161731; DOI=10.1128/JVI.78.12.6381-6388.2004;		
RA	Gao G., Vandenberghe L.H., Alviria M.R., Lu Y., Calcedo R., Zhou X.,		
RT	Wilson J.M.;		
RT	"Clades of Adeno-associated viruses are widely disseminated in human		
RL	tissues.";		
CC	J. Virol. 78:6381-6388(2004).		
CC	Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms		
CC	Distributed under the Creative Commons Attribution-NonCommercial		
CC	License		
DR	EMBL; AY350591; AAC99276.1; -; Genomic_DNA.		
DR	SMR; 06JC28; 217-735.		
DR	GO; GO:0019028; C:viral capsid; IEA.		
DR	GO; GO:0005198; F:structural molecule activity; IEA.		
DR	InterPro; IPR01403; Parvo_coat.		
DR	Pfam; PF00740; Parvo_coat; 1.		
DR	SEQUENCE 735 AA; 81898 MW; 197066F8911FAE9B CRC64;		
QY	Query Match	86.6%; Score 3453.5; DB 2; Length 735;	
	Best Local Similarity	84.5%; Pred. No. 3.2e-216;	
	Matches 622; Conservative 48; Mismatches 65; Indels 1; Gaps 1		
QY	1	MAADGYLEDMLEEDINSEGIREFMMDLKPAPRPKANQOQODDGRGLVLPGYKYLGPFGNGLD	60
DB	1	MAADGYLEDMLEEDLSEGIROMWMLKRPGRPPPKPAERIKDGRGLVLPGYKYLGPFGNGLD	60
QY	61	KGEPVNAADAALAEHDKAYDQOLAKGDNPLYRYNHDAEFOERLOEDTSPFGNIGRAVFO	120
DB	61	KGEPVNEADAALAEHDKAYDRQLNSGDNPLYLYNHADAEPORLEKEDTSPFGNIGRAVFO	120
QY	121	AKKVLVEELGIVEEGAKTAPPEKRPVRESPOEPSSSGTIGTKGOOPAKKRLNFGQYQISE	180
DB	121	AKKVLVEELGIVEEPVKTAPPEKRPVRESPAEPSSSGTIGAGQOPAKKRLNFGQYQISD	180
QY	181	SVPDPQPLGEPAPPAAPGPTTMASSGGGAPADNNEGADGVGNASGNHCDSTWLGDRVI	240
DB	181	SVPDPQPLGEPAPPAAPGSGGTTTMAATGSAAPADNNEGADGVGNSSGNHCDSTWLGDRVI	240
QY	241	TTSTRTTMAALPYNNHLYKOISSASTGASNDNNHYGYSTPWGYPDPNRPFGHSPDPWQRL	300
DB	241	TTSTRTTMAALPYNNHLYKOISSOS-GASNDNNHYGYSTPWGYPDPNRPFGHSPDPWQRL	299
QY	301	INNNMGFPKPKLNFQVKEVTTNDGVTIANNLTSTVOVPSDSYOLPYVLGSAHQ	360
DB	300	INNNMGFPKPKLNFQVKEVTTNDGVTIANNLTSTVOVPTDSYOLPYVLGSAHQ	359
QY	361	GCLPFPADVEMIQYGLTLNNGSOAVGRSSFYCLEYFPQOMLRGNNFTFSYTFEEVP	420

```

Db      360 GCLPPFPADVFMVPOYGYLTLLNGSOAVGRSPFYCLEYFPSPQMLRTGNNFQSFTEBDV 419
QY      421 FHSSVYAHQSGLDRLMNPILIDQYLYLNRTQNGSQAQNKDLFSRSGSPAGMSVQPKXWLP 480
Db      420 FHSSVYAHQSGLDRLMNPILIDQYLYLNRTQNGSQAQNKDLFSRSGSPAGMSVQPKXWLP 479
QY      481 GPCYRQORVSKTKTDNNNSNFTWTGASKYNLNGRESIINPGTAMASHKDEDEKFFPMG 540
Db      480 GPCYRQORVSKTKTDNNNSNFTWTGASKYNLNGRESIINPGTAMASHKDEDEKFFPMG 539
QY      541 MIEGKSAGASNTALDNVMTTDEBEIKATNPVATERFGTVAVNFQSSSTDPADGVYAMG 600
Db      540 LIFGKQGTANADLDLNMVITDEBEIKRTNPVATEQGYTSNNLQNSNTGPTTGTVNHG 599
QY      601 ALPGMWQDRDVLVYLOQPIWAKIPHTDGHFHPSPIMGFGGLKPNPQIILKNTVPANPA 660
Db      600 ALPGMWQDRDVLVYLOQPIWAKIPHTDGHFHPSPIMGFGGLKPNPQIILKNTVPANPA 659
QY      661 EFESATKFASTTQYSTGVSVIEIEMELQKNSKRWNPVQYTSNVAKSANVDFTVNNG 720
Db      660 NFSSAKFASFTTQYSTGVSVIEIEMELQKNSKRWNPVQYTSNVAKSANVDFTVNNG 719
QY      721 YTEPRPIGTGYLTRPL 736
Db      720 YSEPRPIGTGYLTRNL 735

RESULT 15
ID Q6JC19_9VIRU PRELIMINARY; PRT; 738 AA.
AC Q6JC19_
DT 05-JUL-2004, integrated into UniProtKB/TrEMBL.
DT 05-JUL-2004, sequence version 1.
DT 07-FEB-2006, entry version 7.
DE Capsid protein VP1.
GN Name:cap.
OS Adeno-associated virus.
OC Viruses; ssDNA viruses; Parvoviridae; Parvovirinae; Dependovirus;
OC unclassified Dependovirus.
OX NCBI_TaxID=272636;
RN NUCLEOTIDE SEQUENCE.
RP PubMed=15163731; DOI=10.1128/JVI.78.12.6381-6388.2004;
RA Gao G., Vandenberghe L.H., Alvira M.R., Lu Y., Calcedo R., Zhou X.,
RT "Clades of Adeno-associated viruses are widely disseminated in human
RT tissues."
RL J. Virol. 78:6381-6388(2004).
CC -----
CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NonDerivs license
CC -----
DR EMBL; AY530600; AAS9285.1; -; Genomic_DNA.
DR SMR; Q6JC19; 218-738.
DR GO; GO:0019028; C:Viral capsid; IEA.
DR GO; GO:0005198; F:Structural molecule activity; IEA.
DR InterPro; IPR001403; Parvo_coat.
DR Pfam; PF00740; Parvo_coat; 1.
SQ SEQUENCE 738 AA; 81582 MW; 585FA6CF0769934F CRC64;

Query Match      86.6%; Score 3453; DB 2; Length 738;
Best Local Similarity 85.0%; Pred. No. 3.5e-216;
Matches 627; Conservative 48; Mismatches 61; Indels 2; Gaps 2;

```

```

Db      121 AKKRVLEPIGLVEAKTAPGKKRPVPSQSPDSSTGIGKKGGQPAKKRLNFGQTGDS 180
QY      180 ESYVDPQPIGEPPATPAVPTTMAAGGAPMADNNEGADGVGNASGNWCHDSITWIGDRV 239
Db      181 ESYVDPQPIGEPPAGSGGSGTMAAGGAPMADNNEGADGVGSSSGNWHCHDSITWIGDRV 240
QY      240 ITTSTRTMALPTYNHLVYQISASGTAS -NDHYRGSTPMGYPDPNRFCHFSPPDMQ 298
Db      241 ITTSTRTMALPTYNHLVYQISGISTNDTNTTGYSTPMGYFPNRFCHFSPPDMQ 300
QY      299 RLINNMGRPKLNFKNFYQKEVTTNDGVTTLANNLTSTVQVSDSEYQLPYVLGSA 358
Db      301 RLINNMGRPKRLSKLNFNIQKEVTLQNGTKTIANNLSTIQVFTDSEYQLPYVLGSA 360
QY      359 HOGCLPPFPADVFMVPOYGYLTLLNGSOAVGRSPFYCLEYFPSPQMLRTGNNFQSFTEEC 418
Db      361 HOGCLPPFPADVFMVPOYGYLTLLNGSOAVGRSPFYCLEYFPSPQMLRTGNNFQSFTEBD 420
QY      419 VPFSVYAHQSGLDRLMNPILIDQYLYLNRTQNGSQAQNKDLFSRSGSPAGMSVQPKW 478
Db      421 VPFSVYAHQSGLDRLMNPILIDQYLYLSKTSQTGGTQGLLFSQAQPANMSAQAKW 480
QY      479 LPGPCYRQORVSKTKTDNNNSNFTWTGASKYNLNGRESIINPGTAMASHKDEDEKFFPM 538
Db      481 LPGPCYRQORVSKTKTDNNNSNFTWTGASKYNLNGRESIINPGTAMASHKDEDEKFFPM 540
QY      539 GVMIPEGKESAGASNTALDNVMTTDEBEIKATNPVATERFGTVAVNFQSSSTDPADGVYA 598
Db      541 GVLMFGKQAGRDNDVSVMLTSEBEIKTNPVATEQGVADVADNLQQTWTGPIVGNVNS 600
QY      599 MGALPGMWQDRDVLVYLOQPIWAKIPHTDGHFHPSPIMGFGGLKPNPQIILKNTVPANP 658
Db      601 QGALPGMWQDRDVLVYLOQPIWAKIPHTDGHFHPSPIMGFGGLKPNPQIILKNTVPANP 660
QY      659 PAEFSATKFASTTQYSTGVSVIEIEMELQKNSKRWNPVQYTSNVAKSANVDFTVN 718
Db      661 PTFESQAKLASFTTQYSTGVSVIEIEMELQKNSKRWNPVQYTSNVAKSANVDFTVN 720
QY      719 GLYTEPRPIGTGYLTRPL 736
Db      721 GYSEPRPIGTGYLTRNL 738

```

Search completed: June 12, 2006, 12:52:24
 Job time : 308 secs

GenCore version 5.1.9
Copyright (c) 1993 - 2006 Bioacceleration Ltd.

OM protein - protein search, using sw model

Run on: June 12, 2006, 12:52:39 ; Search time 52 Seconds
(without alignments)
1238.895 Million cell updates/sec

Title: US-10-696-261-13

Perfect score: 3989
Sequence: 1 MAADGYLPMLENDLSEGR.....NGLYTPRPITGRITRPL 736

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 650591 seqs, 87530628 residues

Total number of hits satisfying chosen parameters: 650591

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents AA:*

- 1: /EMC_Celerra_SIDS3/prodata/2/1aa/5.COMB.pep:*
- 2: /EMC_Celerra_SIDS3/prodata/2/1aa/6.COMB.pep:*
- 3: /EMC_Celerra_SIDS3/prodata/2/1aa/7.COMB.pep:*
- 4: /EMC_Celerra_SIDS3/prodata/2/1aa/H.COMB.pep:*
- 5: /EMC_Celerra_SIDS3/prodata/2/1aa/PCTUS.COMB.pep:*
- 6: /EMC_Celerra_SIDS3/prodata/2/1aa/RR.COMB.pep:*
- 7: /EMC_Celerra_SIDS3/prodata/2/1aa/backfile1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	3989	100.0	736	2 US-09-807-802A-3	Sequence 3, Appl1
2	3989	100.0	736	2 US-09-807-802A-13	Sequence 13, Appl1
3	3402.5	85.3	735	2 US-09-321-589-1	Sequence 1, Appl1
4	3402.5	85.3	735	2 US-10-293-478-1	Sequence 1, Appl1
5	3402.5	85.3	735	2 US-10-038-972A-13	Sequence 13, Appl1
6	3402.5	85.3	735	2 US-09-717-789C-24	Sequence 24, Appl1
7	3251	81.5	534	2 US-09-807-802A-15	Sequence 15, Appl1
8	2906	72.9	534	2 US-09-807-802A-17	Sequence 17, Appl1
9	2759.5	69.2	534	2 US-10-038-972A-14	Sequence 14, Appl1
10	2486.5	62.3	734	2 US-09-532-594B-4	Sequence 4, Appl1
11	2481.5	62.2	533	2 US-10-038-972A-15	Sequence 15, Appl1
12	2220	55.7	724	2 US-09-533-427-4	Sequence 4, Appl1
13	2220	55.7	724	2 US-09-717-789C-4	Sequence 4, Appl1
14	1830.5	45.9	588	2 US-09-532-594B-16	Sequence 16, Appl1
15	1700.5	42.6	588	2 US-09-533-427-5	Sequence 5, Appl1
16	1700.5	42.6	588	2 US-09-717-789C-5	Sequence 5, Appl1
17	1690.5	42.4	544	2 US-09-532-594B-18	Sequence 18, Appl1
18	1665	41.7	532	2 US-09-533-427-6	Sequence 6, Appl1
19	1665	41.7	532	2 US-09-717-789C-6	Sequence 6, Appl1
20	1410	35.3	756	2 US-09-438-268-4	Sequence 4, Appl1
21	599.5	15.0	781	2 US-10-187-253E-27	Sequence 27, Appl1
22	590.5	14.8	781	2 US-10-187-253E-33	Sequence 33, Appl1
23	486	12.2	554	2 US-10-187-253E-29	Sequence 29, Appl1
24	486	12.2	554	2 US-10-187-253E-35	Sequence 35, Appl1
25	479.5	12.0	543	2 US-08-856-841-22	Sequence 22, Appl1
26	430.5	10.8	500	2 US-08-856-841-16	Sequence 16, Appl1

27	430.5	10.8	501	2 US-08-856-841-18	Sequence 18, Appl1
28	429.5	10.8	486	2 US-08-856-841-19	Sequence 19, Appl1
29	326	8.2	415	2 US-08-856-841-20	Sequence 20, Appl1
30	318	8.0	395	2 US-08-856-841-13	Sequence 13, Appl1
31	310	7.8	264	2 US-08-856-841-14	Sequence 14, Appl1
32	303	7.6	398	2 US-08-856-841-17	Sequence 17, Appl1
33	284	7.1	387	2 US-08-856-841-17	Sequence 17, Appl1
34	216.5	5.4	579	2 5223424-13	Patent No. 5223424
35	192	4.8	584	2 US-09-022-949-2	Sequence 2, Appl1
36	142	3.6	210	2 US-08-856-841-9	Sequence 9, Appl1
37	142	3.6	227	2 US-08-856-841-15	Sequence 15, Appl1
38	142	3.6	489	2 US-08-856-841-12	Sequence 12, Appl1
39	132.5	3.3	489	2 US-10-376-387B-4	Sequence 4, Appl1
40	132.5	3.3	489	2 US-10-889-526-4	Sequence 4, Appl1
41	132.5	3.3	489	2 US-10-889-527-4	Sequence 4, Appl1
42	127	3.2	1287	2 US-10-094-749-2627	Sequence 2627, Ap
43	124	3.1	3060	1 US-08-487-826B-14	Sequence 14, Appl1
44	122.5	3.1	1637	2 US-10-172-502-14	Sequence 14, Appl1
45	121	3.0	2736	2 US-09-252-991A-30227	Sequence 30227, A

ALIGNMENTS

```
RESULT 1
US-09-807-802A-3
; Sequence 3, Application US/09807802A
; Patent No. 6759237
; GENERAL INFORMATION:
; APPLICANT: Wilson, James M.
; TITLE OF INVENTION: Adeno-Associated Virus Serotype I Nucleic Acid Sequences,
; FILE REFERENCE: GNVPM.031USA
; CURRENT APPLICATION NUMBER: US/09/807, 802A
; PRIOR FILING DATE: 2002-02-21
; PRIOR APPLICATION NUMBER: US 60/107, 114
; PRIOR FILING DATE: 1998-11-05
; PRIOR APPLICATION NUMBER: PCT/US99/25694
; PRIOR FILING DATE: 1999-11-02
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 3
; LENGTH: 736
; TYPE: PRT
; ORGANISM: AAV-1
US-09-807-802A-3

Query Match      100.0%; Score 3989; DB 2; Length 736;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 736; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MAADGYLPMLENDLSEGRWMDLKGAKRPYANQKODGGLVPGYKGLGPPNGLD 60
    1 MAADGYLPMLENDLSEGRWMDLKGAKRPYANQKODGGLVPGYKGLGPPNGLD 60
Db 1 MAADGYLPMLENDLSEGRWMDLKGAKRPYANQKODGGLVPGYKGLGPPNGLD 60

QY 61 KGPVNAADAALAEHDAVYQOLKAGDNPYLRYNHADAFOERLOEDTSGCNGRAVFQ 120
    61 KGPVNAADAALAEHDAVYQOLKAGDNPYLRYNHADAFOERLOEDTSGCNGRAVFQ 120
Db 61 KGPVNAADAALAEHDAVYQOLKAGDNPYLRYNHADAFOERLOEDTSGCNGRAVFQ 120

QY 121 AKKRVLEPLGLVEBGAKTAPGKKRPVQOSPOEPDSSGIGKTGOQPAKKRLNFCQTGDS 180
    121 AKKRVLEPLGLVEBGAKTAPGKKRPVQOSPOEPDSSGIGKTGOQPAKKRLNFCQTGDS 180
Db 121 AKKRVLEPLGLVEBGAKTAPGKKRPVQOSPOEPDSSGIGKTGOQPAKKRLNFCQTGDS 180

QY 181 SVDPDPOPLGPPATTPPAVGGTTWASGGGAPMAUNBEGADVGNAAGNWHGDSITLGRVI 240
    181 SVDPDPOPLGPPATTPPAVGGTTWASGGGAPMAUNBEGADVGNAAGNWHGDSITLGRVI 240
Db 181 SVDPDPOPLGPPATTPPAVGGTTWASGGGAPMAUNBEGADVGNAAGNWHGDSITLGRVI 240

QY 241 TTSRTWALPTYNHLYKQISSASTGASNDNHYFGYSTPMPGFDPNRFCHFSFRDQRL 300
    241 TTSRTWALPTYNHLYKQISSASTGASNDNHYFGYSTPMPGFDPNRFCHFSFRDQRL 300
Db 241 TTSRTWALPTYNHLYKQISSASTGASNDNHYFGYSTPMPGFDPNRFCHFSFRDQRL 300

QY 301 INNMGRPKRLNFKLFNIOQKEVTINDGVTTIANLITSTIVQVFSDEYOLPYVLGSAHQ 360
    301 INNMGRPKRLNFKLFNIOQKEVTINDGVTTIANLITSTIVQVFSDEYOLPYVLGSAHQ 360
```

```
Db 301 INNMGFRPRKLNLFKLFNIQKEVTINDGVTIANNLSTVQVFSDEYQLPYVLGSAHQ 360
Qy 361 GCLPPPADVFMIPOYGYLTLNNGSQAVGRSSFYCLEYFP SQMLRTGNFTFSYTEBEVP 420
Db 361 GCLPPPADVFMIPOYGYLTLNNGSQAVGRSSFYCLEYFP SQMLRTGNFTFSYTEBEVP 420
Qy 421 FHSSVYHSOSLDRMLNPLIDQYLYINRTONOGSAQNKDLFSRGS PAGMSVQPKWLP 480
Db 421 FHSSVYHSOSLDRMLNPLIDQYLYINRTONOGSAQNKDLFSRGS PAGMSVQPKWLP 480
Qy 481 GPCYRQORVSKTIDNNNSFTWTGASKYNLNGRESIINPGTAMASHKDEDEKFFPMGSV 540
Db 481 GPCYRQORVSKTIDNNNSFTWTGASKYNLNGRESIINPGTAMASHKDEDEKFFPMGSV 540
Qy 541 MIFGKSAGASNTALDNVMTDEEBEIKATNPVATERFGTVAVNFQSSSTD PATGDVHAMG 600
Db 541 MIFGKSAGASNTALDNVMTDEEBEIKATNPVATERFGTVAVNFQSSSTD PATGDVHAMG 600
Qy 601 ALPGMWQDRDYLQGPIMAKIPIHTDGHFHSPLMGSGFGLKNPPQILLIKNTVPANPPA 660
Db 601 ALPGMWQDRDYLQGPIMAKIPIHTDGHFHSPLMGSGFGLKNPPQILLIKNTVPANPPA 660
Qy 661 EFSATKPFASFTIOYSTGOVSVEIEMELQKENSKRANPEVOYTSNYAKSANVDFTVDNGL 720
Db 661 EFSATKPFASFTIOYSTGOVSVEIEMELQKENSKRANPEVOYTSNYAKSANVDFTVDNGL 720
Qy 721 YTEPRPIGTRYLTRPL 736
Db 721 YTEPRPIGTRYLTRPL 736

RESULT 2
US-09-807-802A-13
; Sequence 13, Application US/09807802A
; Patent No. 6759237
; GENERAL INFORMATION:
; APPLICANT: Wilson, James M.
; APPLICANT: Xiao, Weidong
; TITLE OF INVENTION: Adeno-Associated Virus Serotype 1 Nucleic Acid Sequences,
; TITLE OF INVENTION: Vectors and Host Cells Containing Same
; FILE REFERENCE: GNPV.031USA
; CURRENT APPLICATION NUMBER: US/09/807, 802A
; CURRENT FILING DATE: 2002-02-21
; PRIOR APPLICATION NUMBER: US 60/107,114
; PRIOR FILING DATE: 1998-11-05
; PRIOR APPLICATION NUMBER: PCT/US99/25694
; PRIOR FILING DATE: 1999-11-02
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 13
; LENGTH: 736
; TYPE: PRT
; ORGANISM: AAV-1
US-09-807-802A-13

Query Match 100.0%; Score 3989; DB 2; Length 736;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 736; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MAADGYLPDWLENDLSEGISREWMWLKPGAPKPKANQKODDGRGLVPGYKYLGPFNGLD 60
Db 1 MAADGYLPDWLENDLSEGISREWMWLKPGAPKPKANQKODDGRGLVPGYKYLGPFNGLD 60
Qy 61 KGPENVAAADAALAEHDKAYDQOLKAGDNPLYRYNHADAERFOERLQEDTSGFNLGRAVFQ 120
Db 61 KGPENVAAADAALAEHDKAYDQOLKAGDNPLYRYNHADAERFOERLQEDTSGFNLGRAVFQ 120
Qy 121 AKKRVLEPLGLVEGATAPGKKRPVEQSPQEPDSSSGIGKTGOOPAKKALNFQOTDSE 180
Db 121 AKKRVLEPLGLVEGATAPGKKRPVEQSPQEPDSSSGIGKTGOOPAKKALNFQOTDSE 180
Qy 181 SVDPQGLGEPBPATPAVGPPTTMASSGGAPMAADNNEGADGVGNASGNWHODSTWLGDRVI 240
```

```
Db 181 SVDPQGLGEPBPATPAVGPPTTMASSGGAPMAADNNEGADGVGNASGNWHODSTWLGDRVI 240
Qy 241 TSTRTMALEPTYNHLYKYOISSASTGASNDNHFGYSTWGFYDFENRFHCFSPDMOQL 300
Db 241 TSTRTMALEPTYNHLYKYOISSASTGASNDNHFGYSTWGFYDFENRFHCFSPDMOQL 300
Qy 301 INNMGFRPRKLNLFKLFNIQKEVTINDGVTIANNLSTVQVFSDEYQLPYVLGSAHQ 360
Db 301 INNMGFRPRKLNLFKLFNIQKEVTINDGVTIANNLSTVQVFSDEYQLPYVLGSAHQ 360
Qy 361 GCLPPPADVFMIPOYGYLTLNNGSQAVGRSSFYCLEYFP SQMLRTGNFTFSYTEBEVP 420
Db 361 GCLPPPADVFMIPOYGYLTLNNGSQAVGRSSFYCLEYFP SQMLRTGNFTFSYTEBEVP 420
Qy 421 FHSSVYHSOSLDRMLNPLIDQYLYINRTONOGSAQNKDLFSRGS PAGMSVQPKWLP 480
Db 421 FHSSVYHSOSLDRMLNPLIDQYLYINRTONOGSAQNKDLFSRGS PAGMSVQPKWLP 480
Qy 481 GPCYRQORVSKTIDNNNSFTWTGASKYNLNGRESIINPGTAMASHKDEDEKFFPMGSV 540
Db 481 GPCYRQORVSKTIDNNNSFTWTGASKYNLNGRESIINPGTAMASHKDEDEKFFPMGSV 540
Qy 541 MIFGKSAGASNTALDNVMTDEEBEIKATNPVATERFGTVAVNFQSSSTD PATGDVHAMG 600
Db 541 MIFGKSAGASNTALDNVMTDEEBEIKATNPVATERFGTVAVNFQSSSTD PATGDVHAMG 600
Qy 601 ALPGMWQDRDYLQGPIMAKIPIHTDGHFHSPLMGSGFGLKNPPQILLIKNTVPANPPA 660
Db 601 ALPGMWQDRDYLQGPIMAKIPIHTDGHFHSPLMGSGFGLKNPPQILLIKNTVPANPPA 660
Qy 661 EFSATKPFASFTIOYSTGOVSVEIEMELQKENSKRANPEVOYTSNYAKSANVDFTVDNGL 720
Db 661 EFSATKPFASFTIOYSTGOVSVEIEMELQKENSKRANPEVOYTSNYAKSANVDFTVDNGL 720
Qy 721 YTEPRPIGTRYLTRPL 736
Db 721 YTEPRPIGTRYLTRPL 736

RESULT 3
US-09-321-589-1
; Sequence 1, Application US/09321589
; Patent No. 6498244
; GENERAL INFORMATION:
; APPLICANT: PATEL, SALIL D.
; APPLICANT: MCARTHUR, JAMES G.
; TITLE OF INVENTION: ADENO-ASSOCIATED VIRUS CAPSID IMMUNOLOGIC DETERMINANTS
; FILE REFERENCE: 39672
; CURRENT APPLICATION NUMBER: US/09/321,589
; CURRENT FILING DATE: 1999-05-28
; NUMBER OF SEQ ID NOS: 1
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 735
; TYPE: PRT
; ORGANISM: Adeno-associated virus
US-09-321-589-1

Query Match 85.3%; Score 3402.5; DB 2; Length 735;
Best Local Similarity 83.3%; Pred. No. 7,9e-290;
Matches 613; Conservative 51; Mismatches 71; Indels 1; Gaps 1;

Qy 1 MAADGYLPDWLENDLSEGISREWMWLKPGAPKPKANQKODDGRGLVPGYKYLGPFNGLD 60
Db 1 MAADGYLPDWLENDLSEGISREWMWLKPGAPKPKANQKODDGRGLVPGYKYLGPFNGLD 60
Qy 61 KGPENVAAADAALAEHDKAYDQOLKAGDNPLYRYNHADAERFOERLQEDTSGFNLGRAVFQ 120
Db 61 KGPENVAAADAALAEHDKAYDQOLKAGDNPLYRYNHADAERFOERLQEDTSGFNLGRAVFQ 120
Qy 121 AKKRVLEPLGLVEGATAPGKKRPVEQSPQEPDSSSGIGKTGOOPAKKALNFQOTDSE 180
```

Dh 121 AKKRVLEPLGVEEPAVTAQKKRPVHSPVEPSSSGTGAGQGPARKLNFQOTDAD 180
Qy 181 SVDPQPLGEPBPATPAAVGPTTMASSGGAPWADNEGADGVNAGNWHCDSTWLGDRI 240
Dh 181 SVDPQPLGEPBPATPAAVGPTTMASSGGAPWADNEGADGVNAGNWHCDSTWLGDRI 240
Qy 241 TTSRTTALPTNNHLYKQISSASTGASNDNHFGYSTPMGYFPFNRFCHFSRDMQRL 300
Dh 241 TTSRTTALPTNNHLYKQISSASTGASNDNHFGYSTPMGYFPFNRFCHFSRDMQRL 299
Qy 301 INNMGRPRKLNKLFNIQVEVTDGVTIANNLSTVOVSDEYOLPYLGSANQ 360
Dh 300 INNMGRPRKLNKLFNIQVEVTDGVTIANNLSTVOVSDEYOLPYLGSANQ 359
Qy 361 GCLPPFADVFMIPOGYLTLNNGSOAVGRSSFYCLEYFPGOMLRTGNFTFSYTFEEVP 420
Dh 360 GCLPPFADVFMIPOGYLTLNNGSOAVGRSSFYCLEYFPGOMLRTGNFTFSYTFEEVP 419
Qy 421 FHSYAHOSQSLDRMLNPLIDQYLYLNRTONQSGSAONKULLFSRGSBAGMSVOPKMWLP 480
Dh 420 FHSYAHOSQSLDRMLNPLIDQYLYLNRTONQSGSAONKULLFSRGSBAGMSVOPKMWLP 479
Qy 481 GPCYRQORVSKTKDNNNSFTWTGASKYMLNGRESIIINPTAASHKODEKFFPMQSV 540
Dh 480 GPCYRQORVSKTKDNNNSFTWTGASKYMLNGRESIIINPTAASHKODEKFFPMQSV 539
Qy 541 MIFKESAGASNTALDNVMTDEEBEIKATNPVATEREGYAVANFQSSSTPATGDVYAMG 600
Dh 540 LIFKQOSEKTNVDIEKVMITDEEBEIRTPNVAITEQVGSVSTNQRGNRAAATADVYQG 599
Qy 601 ALPGMWQDDVYIQGPIWAKIPHTDGHFHSPLMGFGKKNPPQILLIKTVPANPRA 660
Dh 600 VLPGMWQDDVYIQGPIWAKIPHTDGHFHSPLMGFGKKNPPQILLIKTVPANPRA 659
Qy 661 EFSATKPSFTTOYSTGVSEIEMELOKENSKRANPEVOYTSYAKSANVDFVDDNGL 720
Dh 660 TFSAKPASFTTOYSTGVSEIEMELOKENSKRANPEIQTSTNYSKSVNVDFTVDINGV 719
Qy 721 YTEBRPIGTRYLTRPL 736
Dh 720 YSEBRPIGTRYLTRNL 735
RESULT 4
US-10-293-478-1
; Sequence 1, Application US/10293478
; Patent No. 6733757
; GENERAL INFORMATION:
; APPLICANT: PATEL, SAILI D.
; APPLICANT: MCARTHUR, JAMES G.
; TITLE OF INVENTION: ADEMO-ASSOCIATED VIRUS CAPSID IMMUNOLOGIC DETERMINANTS
; FILE REFERENCE: 39672
; CURRENT APPLICATION NUMBER: US/10/293,478
; PRIOR FILING DATE: 2002-11-14
; PRIOR APPLICATION NUMBER: US/09/321,589
; NUMBER OF SEQ ID NOS: 1
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 735
; TYPE: PRT
; ORGANISM: Adeno-associated virus
US-10-293-478-1
Query Match 85.3%; Score 3402.5; DB 2; Length 735;
Best Local Similarity 83.3%; Pred. No. 7.9e-290;
Matches 613; Conservative 51; Mismatches 71; Indels 1; Gaps 1;
Qy 1 MAADGYLPDMLDNLSEGIREWMDLKGAPKPKANQKODDGRGLVLPGYKYLGPFGNLD 60
Dh 1 MAADGYLPDMLDNLSEGIREWMDLKGAPKPKABHNKODSRGLVLPGYKYLGPFGNLD 60
Qy 61 KGEFVNADAAALEHDAKAYDQQLKAGDNPYILRYNHADAERLQEDTSGGNGIGRAVFO 120

Dh 61 KGEFVNADAAALEHDAKAYDQQLKAGDNPYILRYNHADAERLQEDTSGGNGIGRAVFO 120
Qy 121 AKKRVLEPLGVEEAGTAAQKKRPVHSPVEPSSSGITGQGPARKLNFQOTDSE 180
Dh 121 AKKRVLEPLGVEEPAVTAQKKRPVHSPVEPSSSGTGAGQGPARKLNFQOTDAD 180
Qy 181 SVDPQPLGEPBPATPAAVGPTTMASSGGAPWADNEGADGVNAGNWHCDSTWLGDRI 240
Dh 181 SVDPQPLGEPBPATPAAVGPTTMASSGGAPWADNEGADGVNAGNWHCDSTWLGDRI 240
Qy 241 TTSRTTALPTNNHLYKQISSASTGASNDNHFGYSTPMGYFPFNRFCHFSRDMQRL 300
Dh 241 TTSRTTALPTNNHLYKQISSASTGASNDNHFGYSTPMGYFPFNRFCHFSRDMQRL 299
Qy 301 INNMGRPRKLNKLFNIQVEVTDGVTIANNLSTVOVSDEYOLPYLGSANQ 360
Dh 300 INNMGRPRKLNKLFNIQVEVTDGVTIANNLSTVOVSDEYOLPYLGSANQ 359
Qy 361 GCLPPFADVFMIPOGYLTLNNGSOAVGRSSFYCLEYFPGOMLRTGNFTFSYTFEEVP 420
Dh 360 GCLPPFADVFMIPOGYLTLNNGSOAVGRSSFYCLEYFPGOMLRTGNFTFSYTFEEVP 419
Qy 421 FHSYAHOSQSLDRMLNPLIDQYLYLNRTONQSGSAONKULLFSRGSBAGMSVOPKMWLP 480
Dh 420 FHSYAHOSQSLDRMLNPLIDQYLYLNRTONQSGSAONKULLFSRGSBAGMSVOPKMWLP 479
Qy 481 GPCYRQORVSKTKDNNNSFTWTGASKYMLNGRESIIINPTAASHKODEKFFPMQSV 540
Dh 480 GPCYRQORVSKTKDNNNSFTWTGASKYMLNGRESIIINPTAASHKODEKFFPMQSV 539
Qy 541 MIFKESAGASNTALDNVMTDEEBEIKATNPVATEREGYAVANFQSSSTPATGDVYAMG 600
Dh 540 LIFKQOSEKTNVDIEKVMITDEEBEIRTPNVAITEQVGSVSTNQRGNRAAATADVYQG 599
Qy 601 ALPGMWQDDVYIQGPIWAKIPHTDGHFHSPLMGFGKKNPPQILLIKTVPANPRA 660
Dh 600 VLPGMWQDDVYIQGPIWAKIPHTDGHFHSPLMGFGKKNPPQILLIKTVPANPRA 659
Qy 661 EFSATKPSFTTOYSTGVSEIEMELOKENSKRANPEVOYTSYAKSANVDFVDDNGL 720
Dh 660 TFSAKPASFTTOYSTGVSEIEMELOKENSKRANPEIQTSTNYSKSVNVDFTVDINGV 719
Qy 721 YTEBRPIGTRYLTRPL 736
Dh 720 YSEBRPIGTRYLTRNL 735
RESULT 5
US-10-038-972A-13
; Sequence 13, Application US/10038972A
; Patent No. 6962815
; GENERAL INFORMATION:
; APPLICANT: J. Bartlett
; TITLE OF INVENTION: AAV VECTORS AND METHODS
; FILE REFERENCE: 28335/36960US
; CURRENT APPLICATION NUMBER: US/10/038,972A
; PRIOR FILING DATE: 2002-01-04
; PRIOR APPLICATION NUMBER: US 60/260,124
; PRIOR FILING DATE: 2001-01-05
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 13
; LENGTH: 735
; TYPE: PRT
; ORGANISM: adeno-associated virus 2 VP1 capsid protein
US-10-038-972A-13
Query Match 85.3%; Score 3402.5; DB 2; Length 735;
Best Local Similarity 83.3%; Pred. No. 7.9e-290;
Matches 613; Conservative 51; Mismatches 71; Indels 1; Gaps 1;
Qy 1 MAADGYLPDMLDNLSEGIREWMDLKGAPKPKANQKODDGRGLVLPGYKYLGPFGNLD 60

Db	1	MAADGYLPLWMEBTDTSSEGIHQWKKLPGPPPPRPAERHKKODSRGLVPGYKYLGPFRNGLD	60
Qy	61	KGEPRNAADAALAEHDKAYDOOLKAGDNYYLRYNHDAEFOERLOEDTSPGNIAGRAVQ	120
Db	61	KGEPRNEDAALAEHDKAYDROLDSDNDYLYKTNHDAEFOERLKHEDTSPGNIAGRAVQ	120
Qy	121	AKKRVLEPLGLVEBEGAKTAPGKKRPVEGSPQEPDSSSGIGTKGOOPAKKRINFGOTDSE	180
Db	121	AKKRVLEPLGLVEEBVKTPGKKRPVEHSRVEDESSSGTGKAOQOPARKRLNFGOTGDAD	180
Qy	181	SVPDPQPLGEPAPATAAANGPTTMASSCGAPMADNNEGAGCVGASGNIWHODSTWLGDRVI	240
Db	181	SVPDPQPLGQPPAASPSGLDNTNMTATGSGAPMADNNEGAGCVGANSNGMHCDSTWMGDRVI	240
Qy	241	TTSTRTMLPTYNHNLKYOISSASSTASGADNHHFGYSTPWGYDPFNRFHCFSPRDMQRL	300
Db	241	TTSTRTMLPTYNHNLKYOISSOS-GASDNDHFGYSTPWGYDPFNRFHCFSPRDMQRL	239
Qy	301	INNNGFPRKRLNFKLENIQVKEVTNDGVTYIANNLTSTVOYVDSSEYQLPYVLGSAHQ	360
Db	300	INNNGFPRKRLNFKLENIQVKEVTNDGVTYIANNLTSTVOYFTDSEYQLPYVLGSAHQ	359
Qy	361	GCLPEPPADVFNITPOYGYTLTNNGSGAVGRSSFYCLEYFPSQMLRTGNNTFSTPEEVP	420
Db	360	GCLPEPPADVFNWPOYGYTLTNNGSGAVGRSSFYCLEYFPSQMLRTGNNTFSTPEEDVP	419
Qy	421	FHSSVAHSOSLRLNMPLEIDOLYLYNTNRONGSQAONKDLLFSRSGPAGMSVQPKWLP	480
Db	420	FHSSVAHSOSLRLNMPLEIDOLYLYSKRNTBPGTTQTSLOLQSGOAGASITRQDSRWLP	479
Qy	481	GPCYRQORVSKTKTDNNNSNFTWTGASKYNLNGRESIINPGTAMASHKODEDKFPWGSV	540
Db	480	GPCYRQORVSKTSADNNSSEYSGATKYHLNGRDSLNVNPGPAMASHKODEEKFPPQSGV	539
Qy	541	MFEGESAGASNTALDNWMTTDBEEIKATNPVATEKPGYAVNVFOSSSTDPAQDVHAMG	600
Db	540	LIFGQGSSEKINVDLEKWMITDEEERTJTNPVATEGYGSVSTLQORNRQAATAADVTOG	599
Qy	601	ALPGWMOORDDVYLOGPIWAKIPTHGTHFSPSLMGFGGLKNPPOILLKNTVPANPAP	660
Db	600	VLPGWMOORDDVYLOGPIWAKIPHTDGHFSPSLMGFGGLKNPPOILLKNTVPANPST	658
Qy	661	EFSATKFASTITQYSTGVQSVLEIWELOKENSKRANPEVOYTSNVAKSANVDFTVDNGL	720
Db	660	TFSAKFASTFIQYSTGVQSVLEIWELOKENSKRANPEIYTSNVAKSANVDFTVDTNGV	719
Qy	721	YTERPPIGTRYILTRPL 736	
Db	720	YSEPRPIGTRYILTRNL 735	
RESULT 6			
US-09-717-789C-24			
; Sequence 24, Application US/09717789C			
; Patent No. 6984517			
; GENERAL INFORMATION:			
; APPLICANT: Chiorini, John			
; APPLICANT: Kotlin, Robert M.			
; APPLICANT: Sater, Brian			
; TITLE OF INVENTION: AAVS VECTOR AND USES THEREOF			
; FILE REFERENCE: 14014.0323U3			
; CURRENT APPLICATION NUMBER: US/09/717,789C			
; CURRENT FILING DATE: 2000-11-21			
; PRIOR APPLICATION NUMBER: PCT/US99/11958			
; PRIOR FILING DATE: 1999-05-28			
; PRIOR APPLICATION NUMBER: 60/087,029			
; NUMBER OF SEQ ID NOS: 26			
; SOFTWARE: FastSeq for Windows Version 4.0			
; SEQ ID NO 24			
; LENGTH: 735			
; TYPE: PRN			

Query Match	85.3%	Score 3402.5	DB 2	Length 735
Best Local Similarity	83.3%	Pred. No. 7.9e-290		
Matches	613	Conservative	51	Mismatches 71; Indels 1; Gaps 1;
<p>ORGANISM: Artificial Sequence</p> <p>FEATURE: Description of Artificial Sequence; No. 6984517e =</p> <p>OTHER INFORMATION: synthetic construct</p> <p>US-09-717-789C-24</p>				
Query	1	MAADGVPDWLEEDNLSEGIREFWMDLKPGA	PKPKANOQOKDDGRGLVPGYKYLGPENGID	60
Db	1	MAADGVPDWLEEDLSTSGIRQWKLKRP	PPRPKPARRHDDSDSGVLPGYKYLGPENGID	60
Query	61	KGEFVNADAAALEHDKAYDQOLKAGDNPYL	RYNHDAEFOERLQEDTSGNLRGAVQ	120
Db	61	KGEFVNADAAALEHDKAYDRQDSGDNPLYL	KYNHADAFOERLKEDTSFGNLRGAVQ	120
Query	121	AKKRVLEPLGVGEAGTAPGKKRPVGSQ	HOEDSSGIGKIQOQAKRLNPGQDGE	180
Db	121	AKKRVLEPLGVGEAVTAPGKKRPVHSS	VEPDSSSGIGKIQOQAKRLNPGQDGD	180
Query	181	SVDPQELGEPPTPAAVGPTTMA	SGGGAAPMADNNGAGVGNA	240
Db	181	SVDPQELGEPPTPAVSGTGNTMAT	SGAPMADNNGAGVGSSGNHCHDSITWMDRVI	240
Query	241	TTSTRTVALPTVNNHLKYKQISSAST	QASNDNHFGYSTWGVDFENRPFCHSP	300
Db	241	TTSTRTVALPTVNNHLKYKQISSAST	QASNDNHFGYSTWGVDFENRPFCHSP	300
Query	301	INNNGRRPRKLNPKLFNIOKREVT	NDGTTTANLSTVQVFSQSEYOLPVLGSAHQ	360
Db	301	INNNGRRPRKLNPKLFNIOKREVT	NDGTTTANLSTVQVFTDSEYOLPVLGSAHQ	360
Query	361	GCLPPPADVEMTPOYGYLTLNNG	SOAVGRSFCLEFPSCMLRTGNFTESYT	420
Db	361	GCLPPPADVEMTPOYGYLTLNNG	SOAVGRSFCLEFPSCMLRTGNFTESYT	420
Query	421	PHSSVYASQSLDRMLNPLIDQYL	YYLNRTQNGSQAONKDLFSSRSPAGMS	480
Db	421	PHSSVYASQSLDRMLNPLIDQYL	YYLNRTQNGSQAONKDLFSSRSPAGMS	480
Query	481	GPCRROORVSKTIDNNNSNFT	TGASKYLNLRBESTINPGTMA	540
Db	481	GPCRROORVSKTIDNNNSNFT	TGASKYLNLRBESTINPGTMA	540
Query	541	MIPEKESASNTLADVMITDEBEI	KATPVATERSFGTAAVAFOSST	600
Db	541	MIPEKESASNTLADVMITDEBEI	KATPVATERSFGTAAVAFOSST	600
Query	601	ALPQWQODRDVYLQGPIMAKI	PHTDGHPHSPLMGFGFLKNP	660
Db	601	ALPQWQODRDVYLQGPIMAKI	PHTDGHPHSPLMGFGFLKNP	660
Query	661	EFSATKFASTITQYSTQGV	VEIEBLOKENSGRANPEVQYTSN	720
Db	661	EFSATKFASTITQYSTQGV	VEIEBLOKENSGRANPEVQYTSN	720
Query	721	YTERPPIGTRYLTPPL	736	
Db	721	YTERPPIGTRYLTPPL	736	
<p>RESULT 7</p> <p>US-09-807-802A-15</p> <p>Sequence 15, Application US/09807802A</p> <p>Patent No. 6759237</p> <p>GENERAL INFORMATION:</p> <p>APPLICANT: Wilson, James M.</p> <p>APPLICANT: Xiao, Weidong</p> <p>TITLE OF INVENTION: Adeno-Associated Virus Serotype I Nucleic Acid Sequences</p> <p>TITLE OF INVENTION: Vectors and Host Cells Containing Same</p> <p>FILE REFERENCE: GNPVN.031USA</p>				


```

; CURRENT APPLICATION NUMBER: US/09/807,802A
; CURRENT FILING DATE: 2002-02-21
; PRIOR APPLICATION NUMBER: US 60/107,114
; PRIOR FILING DATE: 1998-11-05
; PRIOR APPLICATION NUMBER: PCT/US99/25694
; PRIOR FILING DATE: 1999-11-02
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 15
; LENGTH: 599
; TYPE: PR
; ORGANISM: AAV-1
US-09-807-802A-15
```

```

Query Match      81.5%; Score 3251; DB 2; Length 599;
Best Local Similarity 100.0%; Pred. No. 1,2e-276;
Matches 599; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```

QY      138  TAPGKRPVEQSPQEPDSSSGIGTGQOPAKKRLFGGTGSESVDPQPLGEPATPAA 197
DB      1    TAPGKRPVEQSPQEPDSSSGIGTGQOPAKKRLFGGTGSESVDPQPLGEPATPAA 60
QY      198  VGPTTMA SGGAPMADNNEGADGVGNASGNWCHDSITWLGDRVITTSRTTALPTYNHLY 257
DB      61  VGPTTMA SGGAPMADNNEGADGVGNASGNWCHDSITWLGDRVITTSRTTALPTYNHLY 120
QY      258  KQISSASTGASNDNHFGYSTPMGYPDENRPHCHSPRDMQRLINNMGFRPKLNFKL 317
DB      121  KQISSASTGASNDNHFGYSTPMGYPDENRPHCHSPRDMQRLINNMGFRPKLNFKL 180
QY      318  NIQKVEYTTNDGVTIANNLSTVQVPSDSYQLPYVLSAHQCLPPPADVFMIPQYG 377
DB      181  NIQKVEYTTNDGVTIANNLSTVQVPSDSYQLPYVLSAHQCLPPPADVFMIPQYG 240
QY      378  YLTNNGSOAVGRSSFYCLEYFSPQMLRTGNNFTFSYFEEVPHSSVAHSQSILDRLMNP 437
DB      241  YLTNNGSOAVGRSSFYCLEYFSPQMLRTGNNFTFSYFEEVPHSSVAHSQSILDRLMNP 300
QY      438  LIDQYLYLNRTONQSGSAQNKDLLFSRGS PAGMSVQPKMWLPQCYRQQRVSKTKTDNN 497
DB      301  LIDQYLYLNRTONQSGSAQNKDLLFSRGS PAGMSVQPKMWLPQCYRQQRVSKTKTDNN 360
QY      498  NSNTWTGASKTNLNGBSIINPGTAMASHKODEDKFFPMGCVMI FGKESAGASNTALDN 557
DB      361  NSNTWTGASKTNLNGBSIINPGTAMASHKODEDKFFPMGCVMI FGKESAGASNTALDN 420
QY      558  VMIDDEEIKATNVATERFGTAVNFOSSSTDPAITGVHAMGALPGVMQDRDVIYQGP 617
DB      421  VMIDDEEIKATNVATERFGTAVNFOSSSTDPAITGVHAMGALPGVMQDRDVIYQGP 480
QY      618  IWAKI PHTDGHFHPSPLMGFGFLKNPPQILIKNTVPANPAEFSATKFASTFOYSTG 677
DB      481  IWAKI PHTDGHFHPSPLMGFGFLKNPPQILIKNTVPANPAEFSATKFASTFOYSTG 540
QY      678  QVSVEIEMELQKENS KRNPEVQYTSNYAKSANVDFTVNNGLYTEPRPIGTRYLTRPL 736
DB      541  QVSVEIEMELQKENS KRNPEVQYTSNYAKSANVDFTVNNGLYTEPRPIGTRYLTRPL 599
```

```

RESULT 8
US-09-807-802A-17
; Sequence 17, Application US/09807802A
; Patent No. 6759237
; GENERAL INFORMATION:
; APPLICANT: Wileon, James M.
; APPLICANT: Xiao, Weidong
; TITLE OF INVENTION: Adeno-Associated Virus Serotype 1 Nucleic Acid Sequences,
; FILE REFERENCE: Vectors and Host Cells Containing Same
; CURRENT APPLICATION NUMBER: US/09/807, 802A
; CURRENT FILING DATE: 2002-02-21
; PRIOR APPLICATION NUMBER: US 60/107,114
; PRIOR FILING DATE: 1998-11-05
```

```

; PRIOR APPLICATION NUMBER: PCT/US99/25694
; PRIOR FILING DATE: 1999-11-02
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 17
; LENGTH: 534
; TYPE: PR
; ORGANISM: AAV-1
US-09-807-802A-17
```

```

Query Match      72.9%; Score 2906; DB 2; Length 534;
Best Local Similarity 100.0%; Pred. No. 2e-246;
Matches 534; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```

QY      203  MASGGGAPMADNNEGADGVGNASGNWCHDSITWLGDRVITTSRTTALPTYNHLYKQISS 262
DB      1    MASGGGAPMADNNEGADGVGNASGNWCHDSITWLGDRVITTSRTTALPTYNHLYKQISS 60
QY      263  ASTGASNDNHFGYSTPMGYPDENRPHCHSPRDMQRLINNMGFRPKLNFKL 322
DB      61  ASTGASNDNHFGYSTPMGYPDENRPHCHSPRDMQRLINNMGFRPKLNFKL 120
QY      323  EVTTNDGVTIANNLSTVQVPSDSYQLPYVLSAHQCLPPPADVFMIPQYGLTILN 382
DB      121  EVTTNDGVTIANNLSTVQVPSDSYQLPYVLSAHQCLPPPADVFMIPQYGLTILN 180
QY      383  NGSQAVGRSSFYCLEYFSPQMLRTGNNFTFSYFEEVPHSSVAHSQSILDRLMNP 442
DB      181  NGSQAVGRSSFYCLEYFSPQMLRTGNNFTFSYFEEVPHSSVAHSQSILDRLMNP 240
QY      443  LYLNRTONQSGSAQNKDLLFSRGS PAGMSVQPKMWLPQCYRQQRVSKTKTDNNSNFT 502
DB      241  LYLNRTONQSGSAQNKDLLFSRGS PAGMSVQPKMWLPQCYRQQRVSKTKTDNNSNFT 300
QY      503  WTGASKTNLNGRESIINPGTAMASHKODEDKFFPMGCVMI FGKESAGASNTALDNVMTD 562
DB      301  WTGASKTNLNGRESIINPGTAMASHKODEDKFFPMGCVMI FGKESAGASNTALDNVMTD 360
QY      563  EEEIKATNPVATERFGTAVNFOSSSTDPAITGVHAMGALPGVMQDRDVIYQGP 622
DB      361  EEEIKATNPVATERFGTAVNFOSSSTDPAITGVHAMGALPGVMQDRDVIYQGP 420
QY      623  PHTDGHFHPSPLMGFGFLKNPPQILIKNTVPANPAEFSATKFASTFOYSTGQVSVE 682
DB      421  PHTDGHFHPSPLMGFGFLKNPPQILIKNTVPANPAEFSATKFASTFOYSTGQVSVE 480
QY      683  IEMELOKENS KRNPEVQYTSNYAKSANVDFTVNNGLYTEPRPIGTRYLTRPL 736
DB      481  IEMELOKENS KRNPEVQYTSNYAKSANVDFTVNNGLYTEPRPIGTRYLTRPL 534
```

```

RESULT 9
US-10-038-972A-14
; Sequence 14, Application US/10038972A
; Patent No. 6962815
; GENERAL INFORMATION:
; APPLICANT: J. Bartlett
; TITLE OF INVENTION: AAV VECTORS AND METHODS
; FILE REFERENCE: 28335/36946US
; CURRENT APPLICATION NUMBER: US/10/038, 972A
; CURRENT FILING DATE: 2002-01-04
; PRIOR APPLICATION NUMBER: US 60/260,124
; PRIOR FILING DATE: 2001-01-05
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 14
; LENGTH: 538
; TYPE: PR
; ORGANISM: adeno-associated virus 2 VP2 capsid protien
US-10-038-972A-14
```

```

Query Match      69.2%; Score 2759.5; DB 2; Length 598;
Best Local Similarity 82.8%; Pred. No. 1.8e-233;
```

```
Matches 495; Conservative 43; Mismatches 59; Indels 1; Gaps 1;
QY 139 APEKKRPVEOSPOEDSSSGIGKTGQOPAKKRLNFGQTDSESVDPPOPLGEPPATPAAY 198
Db 2 APEKKRPVEHSVPEPSSSGTGKAGQQPARKRLNFGQTDGADSVDPDQPLGQPPAASGL 61
QY 199 GPTTASGGGAPADNNEGADGVGNASGNHCHSTWLGDEVITTTSTRTALPTYYNNHLK 258
Db 62 GTTMTATGSGAPADNNEGADGVGNASGNHCHSTWLGDEVITTTSTRTALPTYYNNHLK 121
QY 259 QISSATGASNDNHYGYSTPMGYPFDPNRFCHFSPPDMQRLINNMMGFPPKLNFLKFN 318
Db 122 QISSQS-GASNDNHYGYSTPMGYPFDPNRFCHFSPPDMQRLINNMMGFPPKLNFLKFN 180
QY 319 IQVEVTTNDGVTTIANNLTSTVQVSDSEYQLPYVLGSAHQCLPFPADVFMIPQY 378
Db 181 IQVEVTTNDGVTTIANNLTSTVQVSDSEYQLPYVLGSAHQCLPFPADVFMIPQY 240
QY 379 LTIANGSOAVGRSSFCLEFPQMLRTGNFTSTFEFVPHSSVAHSQSLDRLMNL 438
Db 241 LTIANGSOAVGRSSFCLEFPQMLRTGNFTSTFEFVPHSSVAHSQSLDRLMNL 300
QY 439 IDOYLKNTQONOSGSAQKDLFSRSPAGMSVQPKMLPGCYRQGRVSKTIDNN 498
Db 301 IDOYLKNTQONOSGSAQKDLFSRSPAGMSVQPKMLPGCYRQGRVSKTIDNN 360
QY 499 SNETGASKYKLNGBRSIINPGTAMASHKDEDEKFPFMSGVMIFGESAGASNTALDNV 558
Db 361 SEVSMGATKYHLNGBRSIINPGTAMASHKDEDEKFPFMSGVMIFGESAGASNTALDNV 420
QY 559 MITDEEIKATNPVATRFCTVAVNPFSSSTDPATGVAHMGALPGVWMDRDVYLOGPI 618
Db 421 MITDEEIKATNPVATRFCTVAVNPFSSSTDPATGVAHMGALPGVWMDRDVYLOGPI 480
QY 619 WAKIPHTDGHFHPSPILMGFGGLKNPFPQILIKTTPVPANPAEFSATKFSFTQYSTGQ 678
Db 481 WAKIPHTDGHFHPSPILMGFGGLKNPFPQILIKTTPVPANPAEFSATKFSFTQYSTGQ 540
QY 679 VSVEIEMELQKENSKRNPPEVOYTSNYAKASANDFTVNNGLYTEPRPIGTRYLTRPL 736
Db 541 VSVEIEMELQKENSKRNPPEVOYTSNYAKASANDFTVNNGLYTEPRPIGTRYLTRPL 598

RESULT 10
US-09-532-594B-4
; Sequence 4, Application US/09532594B
; Patent No. 6468524
; GENERAL INFORMATION:
; APPLICANT: Chordin, John A.
; APPLICANT: Kocin, Robert M.
; APPLICANT: Safer, Brian
; APPLICANT: Davidson, Beverly
; TITLE OF INVENTION: AAV4 VECTOR AND USES THEREOF
; FILE REFERENCE: 14014.025202
; CURRENT APPLICATION NUMBER: US/09/532.594B
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 734
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence; No. 6468524e =
; NAME/KEY: misc feature
; OTHER INFORMATION: AAV4 capsid protein VP1
US-09-532-594B-4
Query Match 62.3%; Score 2486.5; DB 2; Length 734;
Best Local Similarity 63.5%; Pred. No. 2.5e-209;
Matches 475; Conservative 81; Mismatches 161; Indels 31; Gaps 10;
```

```
QY 4 DGYLPWLENDLSEGIREWMWLKPGAPKPKANOXKDDRGVLPGYKYLGFENGLDKE 63
Db 3 DGYLPWLENDLSEGIREWMWLKPGAPKPKANOXKDDRGVLPGYKYLGFENGLDKE 62
QY 64 PVNAADAAALHDKAVDOOLKAGDNFYLRYNHADAFOERLOEDTSFGNLAGAVPAOKK 123
Db 63 PVNAADAAALHDKAVDOOLKAGDNFYLRYNHADAFOERLOEDTSFGNLAGAVPAOKK 122
QY 124 RVLEPGLVEBAGKATPKKKRPVEOSPOEDSSSGIGKTGQOPAKKRLN-GQTGSESV 182
Db 123 RVLEPGLVEBAGKATPKKKRPVEOSPOEDSSSGIGKTGQOPAKKRLN-GQTGSESV 182
QY 183 PDEPQIPGEPPATPAAYGPTTMAAGGAPADNNEGADGVGNASGNHCHSTWLGDEVITTT 242
Db 183 PDEPQIPGEPPATPAAYGPTTMAAGGAPADNNEGADGVGNASGNHCHSTWLGDEVITTT 236
QY 243 STRTMTLPTYYNNHLKYQISSASTAGSNDNHYGYSTPMGYPFDPNRFCHFSPPDMQRLIN 302
Db 237 STRTMTLPTYYNNHLKYRIGLGE---SLQSTNYNGFSTPMGYFDPNRFCHFSPPDMQRLIN 292
QY 303 NNMGFPKRLNFKLFINIQVEVTTNDGVTTIANNLTSTVQVSDSEYQLPYVLGSAHQCL 362
Db 293 NNMGFPKRLNFKLFINIQVEVTTNDGVTTIANNLTSTVQVSDSEYQLPYVLGSAHQCL 352
QY 363 LPEFPADVFMIPQYGY---LTIANGSOAVGRSSFCLEFPQMLRTGNFTSTFEFV 419
Db 353 LPEFPADVFMIPQYGYCGLVGTGTSQQQDTRNAFFCLEFPQMLRTGNFTSTFEFV 412
QY 420 PFHSSVAHSQSLDRLMNLIDQYLKNTQONOSGSAQKDLFSRSPAGMSVQPKMLPGCYR 475
Db 413 PFHSSVAHSQSLDRLMNLIDQYLKNTQONOSGSAQKDLFSRSPAGMSVQPKMLPGCYR 469
QY 476 KMLPQPCTRQGRVSKTIDNNNSFTWAGS---KY---NLNGRESIINPGTAMASHK 528
Db 470 KMLPQPCTRQGRVSKTIDNNNSFTWAGS---KY---NLNGRESIINPGTAMASHK 527
QY 529 DDEDEKFPFMSGVMIFGESAGASNTALDNVMITDEEIKATNPVATRFCTVAVNPFSSS 588
Db 528 PADSK-FSNSQLTFAGPKONGTATVPGLTIFSEELATNTATDIDMGNLQGGQSN 586
QY 589 TDEPATGVAHMGALPGVWMDRDVYLOGPIWAKIPHTDGHFHPSPILMGFGGLKNPFPQIL 648
Db 587 NLPTVRLTLALGAVPGVWMDRDVYLOGPIWAKIPHTDGHFHPSPILMGFGGLKNPFPQIL 646
QY 649 IKNTTPVPANPAEFSATKFSFTQYSTGQVSVEIEMELQKENSKRNPPEVOYTSNYAKS 708
Db 647 IKNTTPVPANPAEFSATKFSFTQYSTGQVSVEIEMELQKENSKRNPPEVOYTSNYAKS 706
QY 709 ANVDFTVNNGLYTEPRPIGTRYLTRPL 736
Db 707 NSLLMAPDAAGKYTEPRPIGTRYLTRPL 734

RESULT 11
US-10-038-972A-15
; Sequence 15, Application US/10038972A
; Patent No. 6962815
; GENERAL INFORMATION:
; APPLICANT: J. Bartlett
; TITLE OF INVENTION: AAV VECTORS AND METHODS
; FILE REFERENCE: 28335/36960US
; CURRENT APPLICATION NUMBER: US/10/038.972A
; CURRENT FILING DATE: 2002-01-04
; PRIOR APPLICATION NUMBER: US 60/260,124
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 15
; LENGTH: 533
; TYPE: PRT
; ORGANISM: adeno-associated virus 2 VP3 capsid protein
US-10-038-972A-15
```

Query Match 62.2%; Score 2481.5; DB 2; Length 533;
Best Local Similarity 83.3%; Pred. No. 4.1e-209;
Matches 445; Conservative 37; Mismatches 51; Indels 1; Gaps 1;

QY 203 MASGGA PMADNNGADGVNAGSNMHCDSITWLGDRVITSTRITWALPTYNHLYKOISS 262
DB 1 MATSGAPMADNNGADGVNAGSNMHCDSITWLGDRVITSTRITWALPTYNHLYKOISS 60

QY 263 ASTGASDNHNYFGYSTPWGYDFNRFCHFSPRDMQRLINNNGFPRKLNFKLNIQVK 322
DB 61 OS -GASDNHNYFGYSTPWGYDFNRFCHFSPRDMQRLINNNGFPRKLNFKLNIQVK 119

QY 323 EYTTNDGVTTIANNLTSTVQVFSDEYQLPYVLGSAHQCLPPRPADVFKIPOYGYLTIN 382
DB 120 EYTTNDGVTTIANNLTSTVQVFTDSEYQLPYVLGSAHQCLPPRPADVFKIPOYGYLTIN 179.

QY 383 NGSOAVGRSSFCLEYPPSOMLRTGNNTFSYTFEEVPHSSVYAHOSLDLMLPLIDQY 442
DB 180 NGSOAVGRSSFCLEYPPSOMLRTGNNTFSYTFEDVPHSSVYAHOSLDLMLPLIDQY 239

QY 443 LYYLARTONOGSAQNKDLLFSRGSFAGMSVQPKMWLPFCYRQORVSKITDNNNSNFT 502
DB 240 LYYLARTONOGSAQNKDLLFSRGSFAGMSVQPKMWLPFCYRQORVSKITDNNNSNFT 299

QY 503 WTGASKYNLNGRESIINPTAMASHKDEDEKFFPMGSGVMTFGKESAGASNTALDNMTD 562
DB 300 WTGASKYNLNGRESIINPTAMASHKDEDEKFFPMGSGVMTFGKESAGASNTALDNMTD 359

QY 563 EEEKRNPAVTEFRGYAVVFOSSSTDPAQDVHAMALRGMTWOPBDVYLOQPIYAKI 622
DB 360 EEEKRNPAVTEFRGYAVVFOSSSTDPAQDVHAMALRGMTWOPBDVYLOQPIYAKI 419

QY 623 PHTDGHFSPBLMGFGFKNPPQIILKNTVPANPAPFASATKAFSTTOSTGOVSVE 682
DB 420 PHTDGHFSPBLMGFGFKNPPQIILKNTVPANPAPFASATKAFSTTOSTGOVSVE 479

QY 683 IEMWLOKENSKRANPEVOYTSNYAKSANDVFTVNNGLYTEPRPIGTRYLTRPL 736
DB 480 IEMWLOKENSKRANPEVOYTSNYAKSANDVFTVNNGLYTEPRPIGTRYLTRPL 533

RESULT 12
US-09-533-427-4
; Sequence 4, Application US/09533427
; Patent No. 6855314
; GENERAL INFORMATION:
; APPLICANT: Chiorini, John
; APPLICANT: Kotin, Robert M.
; APPLICANT: Sater, Brian
; APPLICANT: Davidson, Elizabeth
; APPLICANT: Zahner, Joseph
; TITLE OF INVENTION: AAVS VECTOR FOR TRANSDUCING BRAIN CELLS AND LUNG CELLS
; FILE REFERENCE: 14014.032302
; CURRENT APPLICATION NUMBER: US/09/533.427
; NUMBER OF SEQ ID NOS: 23
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 4
; LENGTH: 724
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:/No. 6855314e =
; OTHER INFORMATION: synthetic construct
US-09-533-427-4

Query Match 55.7%; Score 2220; DB 2; Length 724;
Best Local Similarity 58.6%; Pred. No. 6.4e-166;
Matches 431; Conservative 81; Mismatches 198; Indels 26; Gaps 9;

QY 8 PDMLEDLSEGIRBWMWLKPGAPKPKXANQKQDDGRGIVLPGYXLYGPFNLDSGEPPVNA 67
DB 8 PDMLEF-VGBGLRFFLGLGACPPKPKRPNQHQDARGLVLGYNLYLPGNGLDGRGEFVNR 66

QY 68 ADAALAEHDKAYDOQLKAGDNPYLRYNHADAEOERLEOEDTSFGCNIGRAVFOAKRYLE 127
DB 67 ADEVAERHIDISYNNQLEAGDNPYLRYNHADAEOERLEOEDTSFGCNIGRAVFOAKRYLE 126

QY 128 PLGLVERBARKAPGKKRPVEOSPOEPPDSSSISIGTGOOPAKRLINFGTGSSESVPDPQ 187
DB 127 PFGLEVERBARKAPGKKRIDHFPKRCART-----EEDSKP-----STSSDAEAGPSSQ 176

QY 188 LGSEPPATPA-AVGFPTMAAGGAPMADNNGADGVNAGSNMHCDSITWLGDRVITSTRIT 246
DB 177 QLOIPADPASSLGADTMSAGGGGFLGDNNGADGVNAGSNMHCDSITWLGDRVITSTRIT 236

QY 247 WALPTYNHLYKOISSASTGASDNHNYFGYSTPWGYDFNRFCHFSPRDMQRLINNNG 306
DB 237 WLPSPYNNHGYREIKSGSVDSGNANAYGYSTPWGYDFNRFCHFSPRDMQRLINNNG 296

QY 307 FRPKRLNFKLNIQVEYTTNDGVTTIANNLTSTVQVFSDEYQLPYVLGSAHQCLPPF 366
DB 297 FRPKRLNFKLNIQVEYTTNDGVTTIANNLTSTVQVFTDDYQLPYVVGNGTEGCLPAF 356

QY 367 PADVFMIPOYGYLTIN--NGSOAVGRSSFCLEYPPSOMLRTGNNTFSYTFEEVPHSS 424
DB 357 PPQVFTLPQYGYATLNDNTNPTFRKSSFCLEYPPSOMLRTGNNTFSYTFEEVPHSS 416

QY 425 YAHOSLDLMLPLIDQYLYLARTONOGSAQNKDLLFSRGSFAGMSVQPKMWLPFCYRQ 483
DB 417 FAPSQNLFKLANPLVDQYLYLARTONOGSAQNKDLLFSRGSFAGMSVQPKMWLPFCYRQ 469

QY 484 YRQORVSKITDNNNSNFTWTGASKYNLNGRESIINPTAMASHKDEDEKFFPMGSGV 543
DB 470 YRQORVSKITDNNNSNFTWTGASKYNLNGRESIINPTAMASHKDEDEKFFPMGSGV 529

QY 544 GKESAGASNTAL--DNVMTDEBEIKATNPVATERFCTVAVFOSSSTDPAQDVHAMG 600
DB 530 NSQPANPCTTATYLEGNNLITSESETQPVNRVAVNNGQAMTNNOSSTTAPATGYNIQ 589

QY 601 ALPGWQMDRDPVYLOQPIYAKIPHTDGHFSPBLMGFGFKNPPQIILKNTVPANPAP 660
DB 590 IVPGSVMNERDVYLOQPIYAKIPHTGAHFHSPBLMGFGFGFKNPPQIILKNTVPANP 648

QY 661 EFSATKFASTTOSTGOVSVEIEMWLOKENSKRANPEVOYTSNYAKSANDVFTVNNGL 720
DB 649 EFSATKFASTTOSTGOVSVEIEMWLOKENSKRANPEVOYTSNYAKSANDVFTVNNGL 708

QY 721 YTEPRPIGTRYLTRPL 736
DB 709 YTEPRPIGTRYLTRPL 724

RESULT 13
US-09-717-789C-4
; Sequence 4, Application US/09717789C
; Patent No. 6984517
; GENERAL INFORMATION:
; APPLICANT: Chiorini, John
; APPLICANT: Kotin, Robert M.
; APPLICANT: Sater, Brian
; APPLICANT: Davidson, Elizabeth
; APPLICANT: Zahner, Joseph
; TITLE OF INVENTION: AAVS VECTOR AND USES THEREOF
; FILE REFERENCE: 14014.032303
; CURRENT APPLICATION NUMBER: US/09/717.789C
; NUMBER OF SEQ ID NOS: 26
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 724
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:

```
; OTHER INFORMATION: Description of Artificial Sequence; No. 6984517e =  
; OTHER INFORMATION: synthetic construct  
US-09-717-789C-4
```

```
Query Match          55.7%; Score 2220; DB 2; Length 724;  
Best Local Similarity 58.6%; Pred. No. 6,4e-186;  
Matches 431; Conservative 81; Mismatches 198; Indels 26; Gaps 9;
```

```
QY      8 PDMLEDNLSGEGIREMMDLKPFAKPKYANQKQDQDGRGLVPGYKYLGPENGDLKGEPVNA 67  
      8 PDMLEB-VGEGLREPFGLGEPKPKPMQOHOQAOGLVPGYNYLGPGLGRGEPVNR 66  
      68 ADAALAEHOKAYQOOLKAGDNPLYLRVNHADAEFOERLOEPTSGNNGRAVFOAKKRYLE 127  
      67 ADEVAAEHDSIYNEQLSAGDNPLYLRVNHADAEFOERKLADTSFGNNGRAVFOAKKRYLE 126  
QY      128 PLGLVEGAKTAPKGRPVQEOSPOEPPSSSGIKTGQOPAKKRLNFQGTDSSEVPDPQ 187  
      127 PFLGLVEGAKTAPKGRIDHFKRKAR-----BEDSK-----STSSDAEAGPSGQ 176  
QY      188 LGEPPATPA-AVGPTTMAAGGAPMADNNEGADGVGNASGNMHCDSTWGLDRVITTTSTRT 246  
      177 QLOIPAPQAPASISLADITMASAGCGPLGDNNQADGVGNASGDWHCDSTWMDRIVVTXSTRT 236  
QY      247 WALPTNNHLYKQISSASTGASNDNHFGYSTPMGVFDFNRPHCHSPRDMORLINNMG 306  
      237 WVLPTNNHORYREIKSGSDVGSNANAYFGSTPMGFDFNRFHSHSPRMQRLINNYWG 236  
QY      307 FRPKRLNFKLENIQVKEVTNDGVTITANNLSTVQVFSDEYQLPYVLGSAHQGLPRP 366  
      297 FRRLSLRKVIFENIQVKEVTYQDSTTTIANNLSTVQVFTDDYQLPYVVGNGTEGCLPAF 356  
QY      367 PAVYFMIPOYGYTLN--NSQAVGNASSFYCLEYFSPQMLRTGNPFSTFEVPPHSS 424  
      357 PPVFTLPQYGYATLNRDNTENPTERRSPFCLFEPSPKMLRTGNPFSTFEVPPHSS 416  
QY      425 YASOSIDRLMNLIDQYLYLNRTOQSGAQNKDLLFSRGPAGMSVQK-KNWLPGPC 483  
      417 FASOSIDRLMNLIDQYLYLNRTOQSGAQNKDLLFSRGPAGMSVQK-KNWLPGPC 469  
QY      484 YROORSYKTIIDNNNSNFTWTGASKYMLNGRESIINPGTAMASHKDEDEKFFPMGVMIF 543  
      470 GRQGMNLGSGVNRASVASAFATNRMELGASVQVPPQPMGNTNQLQSGNTVLALEMWIF 529  
QY      544 GKESAGASNTAL--DNWMTDEBEIKATNPVATERGTVAVNFQSSSTDPATGVDHAMG 600  
      530 NSQAPAMPGTATYLEGMMLITSSEETOPVNRVAVNMGQWATNQSSTTAPATGYMLQE 589  
QY      601 ALRGMWQDRDYLQGPIMAKIPTDGHFHPSPLMGGFGLKNPPOLLIKNTPPANPPA 660  
      590 IYQGSVWMEKDVYLGQPIWAKIPETGAHFPSPAMGGFGLKNPPOLLIKNTPPANPPA 648  
QY      661 EFSATKFASTFYQSTGVSEIEMELQKENSRRNNEPVQYTSNAYKASANDFTVDNNGL 720  
      649 SFSQDVAVSSFTYQSTGVSEIEMELQKENSRRNNEPVQYTSNAYKASANDFTVDNNGL 708  
QY      721 YTEPRPIGTTRYLTRPL 736  
      709 YRTTRPIGTTRYLTRPL 724  
Db
```

```
RESULT 14  
US-09-532-594B-16  
; Sequence 16, Application US/09532594B  
; Patent No. 6468524  
; GENERAL INFORMATION:  
; APPLICANT: Chorin, John A.  
; APPLICANT: Kotin, Robert M.  
; APPLICANT: Safier, Brian  
; APPLICANT: Davidson, Beverly  
; TITLE OF INVENTION: AAV4 VECTOR AND USES THEREOF  
; FILE REFERENCE: 14014.0252U2  
; CURRENT APPLICATION NUMBER: US/09/532,594B
```

```
; CURRENT FILING DATE: 2000-03-22  
; NUMBER OF SEQ ID NOS: 22  
; SOFTWARE: FaastSeq for Windows Version 4.0  
; SEQ ID NO 16  
; LENGTH: 598  
; TYPE: PRT  
; ORGANISM: Artificial Sequence
```

```
; OTHER INFORMATION: Description of Artificial Sequence; No. 6468524e =  
; OTHER INFORMATION: synthetic construct  
; NAME/KEY: misc feature  
; OTHER INFORMATION: AAV4 capsid protein VP2  
US-09-532-594B-16
```

```
Query Match          45.9%; Score 1830.5; DB 2; Length 598;  
Best Local Similarity 57.8%; Pred. No. 8e-152;  
Matches 355; Conservative 74; Mismatches 154; Indels 31; Gaps 10;
```

```
QY      138 TAPGKRPVQEOSPOEPPSSSGIKTGQOPAKKRLNF-GQGTDSSEVPDPQLGEPATPA 196  
      1 TAPGKRPVQEOSPOEPPSSSGIKTGQOPAKKRLNF-GQGTDSSEVPDPQLGEPATPA 196  
QY      197 AVGPPTMAAGGAPMADNNEGADGVGNASGNMHCDSTWGLDRVITTTSTRTWALPTNNH 256  
      55 MSDSEMRRAAAGAAVEGGADGVGNASGDWHCDSTWSEGHVTTSTRTWALPTNNH 114  
QY      257 YKQISSASTGASNDNHFGYSTPMGVFDFNRPHCHSPRDMORLINNMGFRPKRLNFKL 316  
      115 YKRLG---SLOSNTYNGSTPMGVFDFNRPHCHSPRDMORLINNMGFRPKRLNFKL 170  
QY      317 FNIQVKEVTNDGVTITANNLSTVQVFSDEYQLPYVLGSAHQGLPRPADVFMIPQY 376  
      171 FNIQVKEVTNSNBEITVANNLSTVQIFADSSYLELYVMDAGGEGSLPEPPNVFVMPQY 230  
QY      377 GY---LTLNNGSQAVERSSFYCLEYFSPQMLRTGNPFSTFEVPPHSSYHQSGLDR 433  
      231 GYGLVGTGNTSQOQTRNMFYCLEYFSPQMLRTGNPFSTFEVPPHSSYHQSGLDR 230  
QY      434 LAMPILIDQYLYLNRQON---QSGAQNKDLLFSRGPAGMSVQK-KNWLPGPC 489  
      291 LAMPILIDQYLYLNRQON---QSGAQNKDLLFSRGPAGMSVQK-KNWLPGPC 347  
QY      490 SKTKTNNNSNFTWTGAS--KY-----NLNGRESIINPGTAMASHKDEDEKFFPMGVMIF 542  
      348 SKTA--NQYKIKATGSDSLIKETHTSLDGRSALTPGPMATAGADSK-FSNGLIF 404  
QY      543 GKESAGASNTALDNWMTDEBEIKATNPVATERGTVAVNFQSSSTDPATGVDHAMGAL 602  
      405 AGEKQNGNTATVPGTLIFTSSEELAATNATDTDMGMLPGDQNSNLPVVDRLTALGAV 464  
QY      603 PGWVWQDRDYLQGPIMAKIPTDGHFHPSPLMGGFGLKNPPOLLIKNTPPANPPA 662  
      465 PGWVWQDRDYLQGPIMAKIPTDGHFHPSPLMGGFGLKNPPOLLIKNTPPANPPA 648  
QY      663 SATKFASTFYQSTGVSEIEMELQKENSRRNNEPVQYTSNAYKASANDFTVDNNGL 722  
      525 SSTPNSFTYQSTGVSEIEMELQKENSRRNNEPVQYTSNAYKASANDFTVDNNGL 584  
QY      723 EPRPIGTTRYLTRPL 736  
      585 EPRPIGTTRYLTRPL 598  
Db
```

```
RESULT 15  
US-09-533-427-5  
; Sequence 5, Application US/09533427  
; Patent No. 6855314  
; GENERAL INFORMATION:  
; APPLICANT: Chorini, John  
; APPLICANT: Kotin, Robert M.  
; APPLICANT: Safier, Brian  
; APPLICANT: Davidson, Elizabeth  
; APPLICANT: Zabner, Joseph
```

```

; TITLE OF INVENTION: AAVS VECTOR FOR TRANSDUCING BRAIN CELLS AND LUNG CELLS
; FILE REFERENCE: 14014.0323U2
; CURRENT APPLICATION NUMBER: US/09/533.427
; CURRENT FILING DATE: 2000-03-22
; NUMBER OF SEQ ID NOS: 23
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 5
; LENGTH: 588
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:/No. 6855314e =
; OTHER INFORMATION: Synthetic construct
; US-09-533-427-5

```

```

Query Match          42.6%; Score 1700.5; DB 2; Length 588;
Best Local Similarity 55.0%; Pred. No. 2,1e-140;
Matches 332; Conservative 64; Mismatches 173; Indels 35; Gaps 7;

```

```

QY      140  PGKKRPVEQSPQEPDSSGIGKTGQPPAKKRLNFGQTGDSSESVPDPQPLGEPPTPA-AV 198
      13  PKRKARTEDSKRSTSS-----DAAGPSGSQQQLDIPAQPASL 52
QY      199  GPTTMASGGAPMADNNEGADGVGNASGNHCDSTWLGDRYITSTRTWALPTNNHLXK 258
      53  GADTMSAGGGGFLGDNNQAGDGVGNASGDWDCSTWMDRVTXSTRTWVLPSTNNHQYR 112
QY      259  QISSASTGASNDNHFGSTPMGYPDFNRFCHFSPRDMORLNNNGFRPKRLNFKLFN 318
      113  EIKSGSYDGSANAAHFGYSTPMGTFDNRFRSHWSPRDMQRLNNYNGFRPSRLVKIFN 172
QY      319  IQVEVTNDGVTIANNLSTVOVFSDESEYQPLVLSAHQGLPPPADVEMIPOYGY 378
      173  IQVEVTVOQDSTTTIANNLSTVOVFTDDYQPLVYVGNGBGCLPAFPQVFTLPQYGY 232
QY      379  LTLN--NGSOAVGRSSFYCLEYFSPQMLRTGNNTFSYTPREVPFHSSYAHQSGLRLN 436
      233  ATLNRDNTENPTERRSFFCLEYFSPQMLRTGNNTFSYTPREVEVPFHSSFAQSQNLFLAN 292
QY      437  PLIDQYLXNLRNTOGSAQNKDILFSRGSAGMSVOP--KNWLPGPCYRQORVSKTKTD 495
      293  PLVDQYLYRFVSTNTGTGVQENKUL-----AGRYANTYKNWFPGMGRTOGNLGSV 345
QY      496  NNNSNFTWTGASKYNLNGRESIINPGTAMASHKDEDEKFFPMGSMVIFGKESAGASNTAL 555
      346  NRASVSFAFATTNRMELEGASYOVPPQPNGMTNNLQGSNTYALENTWIFNSQPANPGTTAT 405
QY      556  ---DNWMTDBEELIKANPVATERFGTVAVNFOSSSTDPRATGVDHAMGALPGMYQDRDV 612
      406  YLEGNMLITSESETQPNRYAVNVGQWATNOSTTAPATGYNLQEIIVPGSYMWERDV 465
QY      613  YLOGPYIAKIPHTDGHFPPSLMGFGFKNPPQIILIKNTPVAPNPPAEFSATKFAFIT 672
      466  YLOGPIYAKIPETICAHFHPSPAMGFGFKNPPMMLIKNTPVPGN-ITSSDVPVSSFIT 524
QY      673  QYSTGQVSEIEMELQENSKRNNEVOYTSNYAKSANVDFTVDNNGLYTEPRDIGTRYL 732
      525  QYSTGQVTEMEWELKKENSKRNNEIQTNNYNDPQFVDFAPDSTGEYRTTRPIGTRYL 584
QY      733  TRPL 736
      585  TRPL 588
DB

```

Search completed: June 12, 2006, 12:54:12
 Job time : 55 secs

THIS PAGE BLANK (USPTO)